

Commentary

Cite this article: Katakam KK, Sethi NJ, Jakobsen JC, and Glud C. (2019) Selective serotonin reuptake inhibitors are still harmful and ineffective. Responses to the comments by Hieronymus *et al.* *Acta Neuropsychiatrica* 31: 276–284. doi: [10.1017/neu.2019.24](https://doi.org/10.1017/neu.2019.24)

Received: 30 March 2019
Revised: 2 May 2019
Accepted: 16 May 2019
First published online: 24 June 2019

Key words:

antidepressants; clinical trials; depression; serotonin; SSRI

Author for correspondence: Kiran Kumar Katakam, Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen. E-mail: kk@ctu.dk

Selective serotonin reuptake inhibitors are still harmful and ineffective. Responses to the comments by Hieronymus *et al.*

Kiran Kumar Katakam¹, Naqash Javaid Sethi¹, Janus Christian Jakobsen^{1,2,3} and Christian Glud^{1,2}

¹The Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark; ²The Cochrane Hepato-Biliary Group, Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark and ³Department of Cardiology, Holbæk Hospital, Holbæk, Denmark

Abstract

In this response, we address point by point the additional issues raised by Hieronymus *et al.* in their second round of critique of our systematic review on selective serotonin reuptake inhibitors for major depression. We repulse that we are biased or mistaken in any major ways. We acknowledge that we missed a few small, mostly unpublished trials, and we made a few minor errors in our systematic review. However, these omissions and errors neither have any impact on our overall results nor on our conclusions. The critique by Hieronymus *et al.* seems to raise questions about their understanding of the systematic review process, and, on several occasions, they wrongly claimed that we made errors. Our analyses should be impartial and free from any biases or prejudices as we do not have any obligation to support the interests of sponsors or other groups.

In Hieronymus *et al.*'s most recent critique (Hieronymus *et al.*, 2018a), they criticised both our original systematic review (Jakobsen *et al.*, 2017a) and our response (Katakam *et al.*, 2018) to their earlier critique (Hieronymus *et al.*, 2018b). We strongly disagree with their claim that we want to portray the selective serotonin reuptake inhibitors (SSRIs) as ineffective and harmful, and that we distorted and misquoted our own data, both in the *BMC Psychiatry* paper (Jakobsen *et al.*, 2017a) and in the lay media (Jakobsen *et al.*, 2017b; TV 2 (Denmark), 2017). Janus Christian Jakobsen has never claimed that we have shown SSRIs to enhance the risk of suicide in any of his appearances in media as alleged by the critiques (Hieronymus *et al.*, 2018a). In fact, we think Hieronymus *et al.* misrepresented his statements and dedicated three paragraphs to criticise us (Hieronymus *et al.*, 2018a). Janus Christian Jakobsen correctly claimed that we have shown SSRIs to enhance the risk of serious adverse events (SAEs) (Jakobsen *et al.*, 2017a). He mentioned the terms suicide and death while giving examples of SAEs in his appearances in Scandinavian media. The direct translation of the sentence in Danish article published in *videnskab.dk* (Jakobsen *et al.*, 2017b) cited by Hieronymus *et al.* regarding SAEs is 'The review shows with great certainty that SSRIs increase the risk of serious adverse events (death, suicide, hospital admission OR any other serious event that is harmful) . . . '.

Hieronymus *et al.* claim that we are reluctant to interpret and report our results in an impartial manner (Hieronymus *et al.*, 2018a). As an example, they cited our statement 'the "true" effect of SSRIs might not even be statistically significant' even though our results showed a highly significant effect ($p < 0.00001$) of SSRIs versus placebo on the Hamilton Depression Rating Scale 17 (HDRS17) (Jakobsen *et al.*, 2017a). We do not agree with this accusation. The above statement is made in the context of observance of a small mean SSRI-placebo difference of approximately two HDRS points. Our results showed that all the trials were at high risk of bias (Jakobsen *et al.*, 2017a). It has repeatedly been shown that trials at high risk of bias tend to overestimate beneficial effects of interventions (Sutton *et al.*, 2000; Hróbjartsson *et al.*, 2012, 2013, 2014; Savović *et al.*, 2012; Lundh *et al.*, 2017). Earlier studies revealed that a number of FDA-registered antidepressant trials with negative results simply reported as 'negative results' never provided actual effect sizes and were never published (Ioannidis, 2008; Turner *et al.*, 2008). Hence, we think that the real 'true' difference between the two intervention groups might be even smaller than observed in our review or in fact non-existent (Jakobsen *et al.*, 2017a). Hence, we justify our statement 'the "true" effect of SSRIs might not even be statistically significant'. We will not be able to assess the 'true' effect of SSRIs before we get adequately conducted and blinded randomised clinical trials comparing SSRIs versus comparable nocebos (active placebo) (Jakobsen *et al.*, 2017a).

Hieronymus *et al.* wonder why we care to present these data and plan to spend time on updating our analysis as all the included trials are at high risk of bias (Hieronymus *et al.*, 2018a).

© Scandinavian College of Neuropsychopharmacology 2019. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.



We think that Hieronymus *et al.* do not show proper understanding of the systematic review process, methodology and principles. We have published a protocol (Jakobsen *et al.*, 2013a) before we started working on the review, where we mentioned how we would proceed with our analyses. Systematic reviews should be updated to include new trials. Hieronymus *et al.* (2018a) cited a recent network meta-analysis on antidepressants by Cipriani *et al.* (2018), where only 9% of the trials were categorised as being at high risk of bias. We clarify that there is a fundamental difference in assessing the overall risk of bias between our review (Jakobsen *et al.*, 2017a) and the review by Cipriani *et al.* (2018). In accordance with the current evidence (Sutton *et al.*, 2000; Hróbjartsson *et al.*, 2012, 2013, 2014; Savović *et al.*, 2012; Lundh *et al.*, 2017), we classified a trial at 'low risk of bias', only if all of the bias domains (generation of allocation sequence, allocation concealment, blinding of study personnel and participants, blinding of outcome assessor, attrition, selective outcome reporting and other bias including sponsorship bias) were classified at 'low risk of bias'. If one or more of the bias domains were classified at 'unclear' or at 'high risk of bias', then the trial was classified at 'high risk of bias'. In contrast, the published protocol (Furukawa *et al.*, 2016) of the Cipriani *et al.* review (Cipriani *et al.*, 2018) states that a trial is classified at low risk of bias if none of the domains described above was rated at high risk of bias and three or less were rated at unclear risk; moderate risk of bias if one was rated at high risk of bias or none was rated at high risk of bias but four or more were rated at unclear risk; and all other trials were assumed at high risk of bias. This Cipriani *et al.* approach is a questionable way of assessing bias risks without support in current evidence (Sutton *et al.*, 2000; Hróbjartsson *et al.*, 2012, 2013, 2014; Savović *et al.*, 2012; Lundh *et al.*, 2017). Moreover, this questionable approach is likely the reason for the fact that we and Cipriani *et al.* reach different assessments. We find our methodology more in line with the results of meta-epidemiological studies (Sutton *et al.*, 2000; Hróbjartsson *et al.*, 2012, 2013, 2014; Savović *et al.*, 2012; Lundh *et al.*, 2017). In fact, there is a lot of criticism (Boesen *et al.*, 2018; Gøtzsche, 2018; Moncrieff, 2018; Timimi *et al.*, 2018; Warren, 2018; Whitaker, 2018) regarding the risk of bias assessment and the conclusions of Cipriani *et al.* network meta-analysis (Cipriani *et al.*, 2018).

Hieronymus *et al.* feel that we should have mentioned in our review that using HDRS17 as a measure of effect markedly underestimates SSRI-induced improvement (Hieronymus *et al.*, 2018a). They claim that the clinical significance of the effect of SSRIs is likely to be considerably higher than the effect captured using HDRS17 (Hieronymus *et al.*, 2018a). However, there is no valid evidence supporting their claim. Moreover, several international institutions recommend HDRS17 for the assessment of depression symptoms and most of the depression trials used HDRS17 in their assessment of efficacy of antidepressants (Committee for Medicinal Products for Human Use, 2013; Sundhedsstyrelsen [The Danish National Board of Health], 2007; Center for Drug Evaluation and Research, 2018). When describing HDRS17 and Montgomery-Asberg Depression Rating Scale (MADRS), the U.S. Food and Drug Administration states on their website that: 'Both scales have undergone a considerable amount of psychometric study and are accepted as valid standards of symptom outcome assessment in studies of major depression' ([https://www.fda.gov/ohrms/dockets/AC/07/briefing/2007-4273b1_04-DescriptionofMADRSHAMDDepressionR\(1\).pdf](https://www.fda.gov/ohrms/dockets/AC/07/briefing/2007-4273b1_04-DescriptionofMADRSHAMDDepressionR(1).pdf)). Even more, the SSRIs received regulatory approvals based on the trial results based on HDRS17. Finally, it must be noted that when assessing the effects of SSRIs using other rating

scales [e.g. MADRS or Beck Depression Inventory (BDI)] the results correspond to the HDRS results, that is, very minimal (or non-existing) beneficial effects (Jakobsen *et al.*, 2017a).

Hieronymus *et al.* cited some examples (see below) of reasons why 'observation should be interpreted with caution' in our review (Hieronymus *et al.*, 2018a). Our point to point responses to each of the examples are below:

- (i) *that we have limited insight into the actual clinical impact of the SAEs tentatively associated with SSRI treatment.* Hieronymus *et al.* should consult the International Committee on Harmonization-Good Clinical Practice (ICH-GCP) guidelines (ICH-GCP, 1996) which all clinical trials ought to follow. This internationally accepted guideline clearly states that it is mandatory to consider all the SAEs occurring in a trial whether the events are associated with the treatment or not. It is always difficult to assess whether a given event is caused by the intervention. For example, a traffic accident might be caused by some of the several adverse effects SSRIs lead to.
- (ii) *that our decisions on whether a certain adverse event should be categorised as serious or not were somewhat arbitrary.* We strongly disagree. We clearly predefined (and used) the GCP definition (Jakobsen *et al.*, 2017a). Nevertheless, the reporting of SAEs in most of the publications was very poor and incomplete. Therefore, we considered some events [e.g. Adamson *et al.* (2015), Claghorn *et al.* (1996)] as SAEs if they met the definition of the SAE according to ICH-GCP guidelines (ICH-GCP, 1996).
- (iii) *our decisions regarding which treatment groups to include regarding Adamson *et al.* (2015) and Pettinati *et al.* (2010) and the extent of follow-up phases.* We explicitly included trials comparing SSRIs versus no intervention, placebo, or 'active' placebo in our review (Jakobsen *et al.*, 2013a). In the Adamson *et al.* trial (Adamson *et al.*, 2015), there are two intervention groups: citalopram and placebo. Naltrexone was prescribed for both the intervention groups. Hence, we believe we did not make any mistake in including this trial. Please see below regarding the Pettinati *et al.* trial (Pettinati *et al.*, 2010). Regarding the extent of follow-up phases, we planned to report results assessed both at end of treatment and at maximum follow-up. But due to very limited data at maximum follow-up (making selection bias likely), we only reported results at end of treatment (Jakobsen *et al.*, 2017a).
- (iv) *that trial reports often detail potential SAEs in the active treatment group (this being the issue of interest) while not providing corresponding information regarding similar events in patients on placebo* (Wernicke *et al.*, 1988; Claghorn *et al.*, 1996; Feighner & Overo, 1999). Hieronymus *et al.* do not refer to any valid evidence supporting their claim, perhaps because the evidence does not exist. We do not agree that trial reports often detail potential SAEs in the active treatment group, and we wonder why Hieronymus *et al.* came to such a conclusion? One cannot evaluate events in an active intervention group without having similar unbiased assessments in the control groups.
- (v) *that our SAE data from older trials must generally be interpreted with caution.* We wonder why Hieronymus *et al.* think that older trials are different from newer trials regarding SAE data. In fact, many studies reveal that the reporting of adverse event data is still suboptimal even after introduction of some standards (e.g. CONSORT) for reporting of adverse events

Table 1. Summary of our results of selective reuptake inhibitors versus placebo or no intervention on serious adverse events before (Jakobsen *et al.*, 2017a) and when the valid issues raised by Hieronymus *et al.* (Hieronymus *et al.*, 2018a,b; Katakam *et al.*, 2018) were addressed

Data set	Result on full population odds ratio (95% confidence interval) with <i>p</i> value	<i>p</i> value in elderly subgroup	<i>p</i> value in non-elderly subgroup
Our original data set (Jakobsen <i>et al.</i> , 2017a)	1.31 (1.04–1.66); <i>p</i> = 0.023	<i>p</i> = 0.006	<i>p</i> = 0.320
Our updated data set (Katakam <i>et al.</i> , 2018) after Hieronymus <i>et al.</i> 's first critique (Hieronymus <i>et al.</i> , 2018b)	1.33 (1.08–1.64); <i>p</i> = 0.008	<i>p</i> = 0.005	<i>p</i> = 0.149
Our presently updated data set after Hieronymus <i>et al.</i> 's second critique (Hieronymus <i>et al.</i> , 2018b) (full data set)	1.29 (1.06–1.58); <i>p</i> = 0.017	<i>p</i> = 0.005	<i>p</i> = 0.250

(Ioannidis, 2009; Haidich *et al.*, 2011; Shukralla *et al.*, 2011; Bagul & Kirkham, 2012; Smith *et al.*, 2012; Hodkinson *et al.*, 2013; Péron *et al.*, 2013).

We do not agree with Hieronymus *et al.*'s claim that we are biased regarding reporting of 'high risk of publication bias' (Hieronymus *et al.*, 2018a). Hieronymus *et al.* cited our exchange with the peer reviewers made public by *BMC Psychiatry*. Here we made a statement that our material was skewed in favour of SSRI-positive studies regarding 'high risk of publication bias' and presented significant outcome of an Egger test to back this up. When we found out that we made an error in our assessment, we withdrew our statement and the outcome of the Egger test in our final version of the manuscript (Jakobsen *et al.*, 2017a). Such revisions are natural parts of any peer review process. We do not understand how it can be considered as bias. Our mistake was unintentional. Do Hieronymus *et al.* want us to make a statement which was found out to be an error during the peer reviewing process?

As explained above, we included trials comparing SSRIs versus no intervention, placebo or 'active' placebo in our review (Jakobsen *et al.*, 2013a). We did not make any mistake in including the Ball *et al.* trial (Ball *et al.*, 2014) with no placebo group, but we made a mistake in the reporting of groups; instead of reporting the aprepitant plus paroxetine group versus the aprepitant group, we reported the aprepitant plus paroxetine group versus the paroxetine group. However, this change does not noticeably change our results or conclusions (Table 1). Regarding the Pettinati *et al.* trial (Pettinati *et al.*, 2010), we acknowledge that we only included two treatment groups (sertraline vs. placebo) but missed two other treatment groups (sertraline plus naltrexone vs. naltrexone). However, the updated analysis does not in any way change our results and our conclusions (Table 1).

Regarding omission of an escitalopram arm in the SCT-MD-01 trial (Forest Laboratories Inc, 2001), we think Hieronymus *et al.* (2018a) did not fully understand our methodology in our systematic review (Jakobsen *et al.*, 2017a) and our explanation in our earlier response (Katakam *et al.*, 2018). As the SCT-MD-01 trial (Forest Laboratories Inc, 2001) is a multi-group trial, we subdivided the trial into three experimental groups [escitalopram 10 mg – SCT-MD-01 (A); escitalopram 20 mg – SCT-MD-01 (B); and citalopram 40 mg – SCT-MD-01 (C)], and subdividing the placebo group into three groups to correspond to each of the experimental SSRI group. As there were only two SAEs in the placebo group, we randomly distributed these events to the SCT-MD-01 (B) and SCT-MD-01 groups (C). As there were no SAEs in both the SSRI group and the corresponding placebo group in comparison A, both were excluded and hence there is no question of inflating the apparent rate of SAEs as claimed by Hieronymus *et al.* (2018a).

Regarding exclusion of female-specific SAEs in the GSK/810 trial (GlaxoSmithKline, 2005e), we do not agree with Hieronymus *et al.*'s claim that we adopted opposite policies when extracting data from a similar GSK trial presenting separate sets of fatal and non-fatal SAEs (GlaxoSmithKline, 2005a,d). In those reports, though fatal and non-fatal SAEs were presented as separate sets, they were reported in the same table and they used double asterisk symbol (**) if different events occurred in the same patient. But in the case of the GSK/810 trial report (GlaxoSmithKline, 2005e), female-specific events were reported in a separate table, and hence it was not clear whether the same participants had any other SAEs that were reported in the main table in that report. Anyhow, there were two female-specific SAEs in the placebo group and one female-specific SAE in the SSRI group. Inclusion of these events did not in any way change our results and conclusions (OR 1.26, 95% confidence interval 1.03–1.53, *p* = 0.026).

Hieronymus *et al.* in their earlier critique (Hieronymus *et al.*, 2018b) criticised that our review (Jakobsen *et al.*, 2017a) was marred by many factual errors and inconsistencies. Though we acknowledged in our earlier response (Katakam *et al.*, 2018) that they have identified some errors, we did not agree with Hieronymus *et al.* regarding several of the 'errors' they claim that we made. In their new critique, Hieronymus *et al.* claim that many errata that they listed in their earlier criticism were just examples, and to be seen as illustrations of a flawed process (Hieronymus *et al.*, 2018a). It is strange and difficult to understand that even after our point to point explanation to their earlier critique, Hieronymus *et al.* stick to their earlier stand and still consider them as errors and describe the process as flawed. To illustrate that there were many errors in the review, Hieronymus *et al.* provided additional examples of mistakes in the Supplementary Material. We do not agree with most of their claims and our detailed explanation can be seen in our responses in the Supplementary Material. We think Hieronymus *et al.* wrongly considered several as errors, for example, Hieronymus *et al.* claim that we used pre-treatment values instead of post-treatment values in the trial by Jindal *et al.* (2003). This trial investigated the impact of sertraline on the sleep of depressed patients. In Table 1 of the manuscript by Jindal *et al.* (2003), they reported baseline parameters, and in Table 2, they reported variables of both pre-sleep and post-sleep data. We used the data for pre-sleep. Hieronymus *et al.* might have confused pre-sleep with pre-treatment. There are several other similar examples (see Supplementary Material).

Hieronymus *et al.* also stressed that they are just illustrative samples that were identified upon their relatively cursory review, and anyone caring to take a closer look would probably find more to add to the errata list (Hieronymus *et al.*, 2018a). We think that even after conducting our review and responses with rigor and impartiality, errors and inaccuracies (e.g. data entry errors and transposition errors) are bound to happen in systematic reviews

and meta-analyses, due to the involvement of large number of people and enormous amount of work load in screening large number of publications and identifying and extracting data from relevant studies. However, we believe that these errors and inaccuracies should not materially affect the overall results and conclusions.

In their earlier critique (Hieronymus *et al.*, 2018b), Hieronymus *et al.* criticised our review (Jakobsen *et al.*, 2017a) of having missed several trials for which SAE data are readily available. In our response (Katakam *et al.*, 2018), we expressed surprise over Hieronymus *et al.*'s conclusion that there was no significant difference between SSRI and placebo with respect to SAEs without including data from the missed trials. In their new critique (Hieronymus *et al.*, 2018a), they justify their action saying that they repeated our analysis using the same trials that were included in our analysis (Jakobsen *et al.*, 2017a) merely to demonstrate the lack of robustness in our results. Their justification is not impressive. We think that they seem more interested in proving our results as not robust than to investigate whether there was an actual association between SSRIs and occurrence of SAEs.

Hieronymus *et al.* claim that we denounced and discarded previous meta-analyses in this field citing 'not searching all relevant databases' as one reason. We do not agree with this claim as we only stated the limitations of previous meta-analyses, and we clarify that we searched all relevant databases. It is inevitable that systematic reviews miss few trials due to a variety of reasons, especially when searching for unpublished reports where it is not possible to perform a systematic search in databases, etc. Hieronymus *et al.* seem to ignore that when our review was published, we included more than twice the number of trials than any other previous meta-analysis or systematic review (Khan *et al.*, 2002; Kirsch *et al.*, 2008; Arroll *et al.*, 2009; Fournier *et al.*, 2010; Gibbons *et al.*, 2012; Undurraga & Baldessarini, 2012).

Hieronymus *et al.* questioned how we managed to locate two Eli Lilly-sponsored studies, HMAQa (Eli Lilly, 2004a) and HMAQb (Eli Lilly, 2004b), and seem to have missed two other studies, HMAQc (Eli Lilly, 2004c) and HMAQd (Eli Lilly, 2004d), in the same repository. We hereby clarify that we searched the repositories of pharmaceutical companies that produce SSRIs (Jakobsen *et al.*, 2013a), and when we searched the Eli Lilly repository with the search term 'fluoxetine' during our earlier search, we did not get any of the above studies. We were able to locate the Eli Lilly-sponsored trials, HMAQb (Eli Lilly, 2004b) and HMAQa (Eli Lilly, 2004a) as we found the published papers (Goldstein *et al.*, 2002, 2004) of these two trials during our screening. We did not search the repositories of the pharmaceutical companies Pharmacia & Upjohn and Novartis as these companies do not produce any of the SSRIs. Hence, we could not identify the three unpublished trials of Pharmacia & Upjohn (2001a,b,c) and an unpublished trial of Novartis (Novartis, 2009). We acknowledge that we missed one GSK trial (GlaxoSmithKline, 2005b) but we do not agree with Hieronymus *et al.* that we have missed another GSK trial (Study No.: 29060/442) (GlaxoSmithKline, 2005f). We excluded this trial as the primary diagnosis is dysthymia concomitant with depression. In our review, we included only trials where the primary diagnosis was major depressive disorder (Jakobsen *et al.*, 2013a). Hieronymus *et al.* claim that while we included one trial of a substance P antagonist that did not include a placebo arm (Ball *et al.*, 2014), we missed four additional studies (reported in two publications) regarding the same drug that were actually both placebo- and paroxetine-controlled (Keller *et al.*, 2006; Liu *et al.*, 2008). During our initial searches, it was not considered possible to go through the full texts of large number of records

obtained. First, we screened the records based on titles and abstracts. But for some records, we did not find abstracts in the database library and we only screened those records based on titles. In such instances, we only screened the full text if the title gave indication that the reported trial was related to SSRI and was a randomised trial. The titles 'Lack of efficacy of the substance P (neurokinin 1 receptor) antagonist aprepitant in the treatment of major depressive disorder' and 'Is bigger better for depression trials?' did not give any indication that they are randomised trials of SSRIs. Therefore, they were excluded. In our review, we only included the trials where the diagnosis of major depressive disorder was made based on one of the standardised criteria, such as ICD 10 (World Health Organization, 1993), DSM III (American Psychiatric Association, 1980), DSM III-R (American Psychiatric Association, 1987), DSM IV (American Psychiatric Association, 1994) or the Feighner criteria (Feighner *et al.*, 1972). Some of the trials which Hieronymus *et al.* claim that we missed were excluded in our review because the reports did not mention how major depressive disorder was diagnosed (Massana, 1998; Eli Lilly, 2014). Hieronymus *et al.* made a mistake in claiming that we missed the Eyding *et al.* study (Eyding *et al.*, 2010), which is actually a systematic review and does not report any trial. Regarding the trials NCT00636246 and NCT00406952 (Pfizer, 2008a,b), results were not reported for these studies at clinicaltrials.gov and there were no records when we searched the repository of Pfizer (https://www.pfizer.com/research/research_clinical_trials/trial_results) with these NCT numbers. As all our searches were conducted before 2016, we could not identify the trial results of protocol CL3-01574-237 which were published in October 2016 (EudraCT, 2016). We acknowledge that we missed one published trial (Gastpar *et al.*, 2006). It is unfortunate that we were unable to locate few trials, and we are thankful to Hieronymus *et al.* for drawing our attention to these missed trials. Nevertheless, we reiterate that Hieronymus *et al.* seem to ignore that when our review was published, we included more than twice the number of trials than any other previous meta-analyses and systematic review (Khan *et al.*, 2002; Kirsch *et al.*, 2008; Arroll *et al.*, 2009; Fournier *et al.*, 2010; Gibbons *et al.*, 2012; Undurraga & Baldessarini, 2012).

We have now included the data from missed trials and performed HDRS17 efficacy analysis. Random-effects meta-analysis of the updated data revealed a mean difference between SSRIs versus no SSRIs of -2.06 points (95% CI -2.36 to -1.75 ; $p < 0.00001$), which is 0.12 HDRS17 points different compared with that of our published review (Jakobsen *et al.*, 2017a). This has no clinical impact on the results or our conclusions of our systematic review.

We do not agree with Hieronymus *et al.* (2018a) when they claim that we confirmed deviating from our protocol in our response (Katakam *et al.*, 2018), and we also do not agree with their statement 'deviating from the protocol is usually regarded as a felony of the gravest kind by Cochranists and treated it as lapse...'. We are in fact surprised by this statement as we have never deviated from the protocol. We have clearly mentioned in our protocol that 'We will undertake this meta-analysis according to the recommendations stated in The Cochrane Handbook for Systematic Reviews of Interventions (Higgins & Green, 2011)'. The Cochrane Handbook mentioned computational problems when no events are observed in one or both groups and suggests alternative non-fixed zero-cell corrections as explored by Sweeting *et al.* (2004). It is clearly stated in the Cochrane Handbook that in case of rare events '... including a correction proportional to the reciprocal of the size of the contrasting study arm, which they

found preferable to the fixed 0.5 correction when arm sizes were not balanced...'. Hence, we do not agree with Hieronymus *et al.*'s claim that we again failed to implement the procedure according to the recommendations of Sweeting *et al.* (2004) with regards to the use of reciprocal zero-cell correction. Hence, we reject the Hieronymus *et al.*'s statement that 'the results would not have been those that Jakobsen and co-workers must have hoped for'. Hieronymus *et al.* expressed surprise that we have refrained from using Sweeting's method for events that were even rarer than SAEs in general, such as individual adverse events including suicides, suicide attempts and suicidal ideation. We clarify that the data for these events were so limited that whatever method we use, there were no enough information to confirm or reject even very large effects of SSRIs.

We followed our protocol in assessing our results regarding the effect of SSRIs on occurrence of SAEs in all trials without accounting for age and reported a *p* value of 0.009 in our systematic review (Jakobsen *et al.*, 2017a). We reanalysed the data and reported a *p* value of 0.002 in our response to Hieronymus *et al.*'s earlier critique (Hieronymus *et al.*, 2018b) after correcting valid mistakes and inclusion of missed trials. An updated reanalysis after inclusion of data from trials that were reported missed by Hieronymus *et al.* in their recent critique (Hieronymus *et al.*, 2018a) still confirms our earlier conclusions. SSRIs significantly increase the risk of an SAE, the *p* value now being 0.012.

Hieronymus *et al.* introduced subgroup analysis of SAE according to age groups in their earlier critique (Hieronymus *et al.*, 2018b). We reported a *p* value of 0.045 for non-elderly group (Katakam *et al.*, 2018). Our updated reanalysis revealed a *p* value of 0.22 in the non-elderly population. However, it must be stressed that such post hoc sensitivity analyses must only be regarded as hypothesis generating and can of course not change the overall results and conclusions! When analysing large data sets such as ours, problems with multiplicity will lead to several random errors caused by multiple comparisons. If Hieronymus *et al.* believe that SSRIs offer more benefit than harm in certain patient groups, then they must present valid evidence confirming this claim.

Hieronymus *et al.* in their recent critique (Hieronymus *et al.*, 2018a) presented results from sensitivity analyses of four overlapping populations, but they excluded six trials which are deemed eligible by us from all their analyses citing different reasons. We do not agree with the exclusion of these trials for the following reasons. As reporting of SAEs was very poor, we considered events that fulfil the definition of SAE according to the ICH-GCP guidelines (ICH-GCP, 1996). In the Adamson *et al.* publication (Adamson *et al.*, 2015), it is mentioned that two patients were unblinded during treatment period for suicidal ideation and severe abdominal cramps. We believe these events are SAEs as it is mentioned in the publication that these events needed unblinding of patients (Adamson *et al.*, 2015). In the Claghorn *et al.* publication (Claghorn *et al.*, 1996), it is mentioned that three fluvoxamine-treated patients had clinically significant electrocardiogram deteriorations which we considered as SAEs. Ravindran *et al.* (1995), under the heading 'Safety', clearly mentioned the word 'serious side effects' referring to Table 2 in their publication; hence we still believe that they are SAEs. Hence, we do not agree with Hieronymus *et al.*'s decision to exclude these three trials owing to 'not presenting SAEs and/or selectively presenting potential SAEs'. Hieronymus *et al.* excluded the Ball *et al.* trial (Ball *et al.*, 2014) for not being placebo-controlled. But it is surprising that they included the comparison of naltrexone versus naltrexone plus sertraline from Pettinati *et al.* trial (Pettinati *et al.*, 2010) in their

recent analyses (Hieronymus *et al.*, 2018a). As explained earlier, we explicitly included trials comparing SSRIs versus no intervention, placebo or 'active' placebo in our review, and hence we do not agree with the exclusion of the Ball *et al.* trial (Ball *et al.*, 2014). Hieronymus *et al.* also excluded one trial for being partially uncontrolled (GlaxoSmithKline, 2005c). Although this exclusion can be debated (there were data on adverse events in the SSRI exposed as well as in placebo control for one centre and active control from another centre), we now also excluded this trial from our reanalysis. For the Mancino *et al.* trial (Mancino *et al.*, 2014), Hieronymus *et al.* claim that there did not occur any SAEs in the relevant arms according to the study report on ClinicalTrials.gov (2011). We do not agree with Hieronymus *et al.* as in Fig. 1 of the publication Mancino *et al.* (2014), it is clearly mentioned that one person in the sertraline group is hospitalised. We think Hieronymus *et al.* are biased in this regard as they selectively picked report on ClinicalTrials.gov (2011) which suited their claim and ignored publication of the trial which reported the event. Though we do not agree with their results due to exclusion of several trials that are deemed eligible by us, it is surprising that Hieronymus *et al.* are still adamant to reconsider the conventional view regarding the tolerability of the SSRIs even after they found a significant difference in the elderly subgroup (*p* range: 0.007–0.011) for all four populations of our data.

We agree with Hieronymus *et al.*, however, that there may still be some trials that have been overlooked as it is not possible to retrieve all the trials that have been conducted. Even if we find all the trials, there is no guarantee that we retrieve all SAE data as only around half the events appear in published journals (Hughes *et al.*, 2014). Systematic reviewing is an ongoing process and updates to a systematic review are conducted at regular intervals to include missed and new trials, and the results are published after each update.

We have now conducted analyses for three different sets of data, namely (i) data from the trials that were included in our original review (Jakobsen *et al.*, 2017a) after correcting an event in placebo group for the trial 99024 (Lundbeck, 2005) and including the SAE events in the comparison of naltrexone versus naltrexone plus sertraline in Pettinati *et al.* trial (Pettinati *et al.*, 2010); (ii) the latter data plus data from the additional trials that were included in our earlier response (Katakam *et al.*, 2018) plus exclusion of one trial (GlaxoSmithKline, 2005c) for being partially uncontrolled; and (iii) the latter data plus data from the trials that were reported missing by Hieronymus *et al.* (2018a) and judged eligible by us (Eli Lilly, 2004c; GlaxoSmithKline, 2005b; Gastpar *et al.*, 2006; Keller *et al.*, 2006; Pfizer, 2008c,d; Novartis, 2009). We did not consider the serious treatment-emergent signs and symptoms as SAEs in three unpublished trials of Pharmacia & Upjohn (2001a,b,c). The reanalysis of the data showed that association between SSRIs and SAEs is still significant for the full population (Table 1) and confirm our results presented in the original review (Jakobsen *et al.*, 2017a). Moreover, the fact that there may be age strata without statistical significance regarding SAEs does not exclude that one should consider using such interventions merely based on the risks of the occurrence of SAEs (European Medicines Agency, 2017).

We acknowledge that Jakobsen and Gluud in an earlier publication regarding HDRS17 concluded that 'There seems to be a need for other more clinically relevant assessment methods' (Jakobsen *et al.*, 2013b). We have never claimed that HDRS17 is the perfect scale. That was also the reason why we planned to include other depression rating scales (e.g. MADRS or BDI) in

our review, and the results when using the other scales were and are very similar compared to the HDRS17 results, that is, effect sizes far below sensible thresholds for clinical significance. We focused on the HDRS17 scale as it is the most widely used depression rating scale and accepted internationally. Moreover, HDRS17 is the recommended depression rating scale used by the international psychiatric society. Hieronymus *et al.* claimed that in our review, we refrained from mentioning that ‘the shortcomings marring the HDRS17 have been suggested to make the difference between active drug and placebo appear smaller than it actually is’ (Hieronymus *et al.*, 2018a). We emphasise that our objective of our systematic review was to assess the beneficial and harmful effects of SSRIs, and it was not to assess the psychometric validity of HDRS17. Furthermore, the results using all other scales show similar results and we have not identified any valid evidence confirming their claim.

We think that the efficacy data presented by our group confirmed previous reports, and we do not agree with Hieronymus *et al.*'s claim that we are mistaken when arguing that these results suggest the effect of SSRIs to be clinically insignificant. We think it is because they do not believe in the use of HDRS17 scale as a measure of effect and suggest using alternative measure like HDRS6. To support their claim, they cited an earlier study (Hieronymus *et al.*, 2016) which is a patient-level post hoc analysis of 18 industry-sponsored placebo-controlled trials of paroxetine, citalopram, sertraline or fluoxetine. The authors reported a standardised mean difference (SMD) effect size of -0.35 when using HDRS6 and -0.27 when using HDRS17 in favour of SSRI (Jakobsen *et al.*, 2013b). Please note that an SMD of 0.35 is far below the National Institute for Clinical Excellence (NICE) threshold for clinical significance (0.5 SMD). Furthermore, if a standard deviation of 10 points is assumed, this difference corresponds to 0.8 HDRS17 points. If Hieronymus *et al.* think that we are mistaken regarding clinical significance, then we wonder how much SMD on the HDRS17 would they consider clinically significant? We used a mean difference of more than 3 points on the HDRS17 as clinically significant (Jakobsen *et al.*, 2013a) as previously recommended by the NICE (Kirsch *et al.*, 2008; Fournier *et al.*, 2010; Mathews *et al.*, 2015). In fact, earlier studies showed that a difference 3 points on the HDRS17 is considered as ‘no clinical change’ and cannot usually be detected by clinicians (Leucht *et al.*, 2013; Moncrieff & Kirsch, 2015). There should be a minimum mean difference of 7 points or more to show a meaningful improvement (Moncrieff & Kirsch, 2015). Hence, we do not agree with Hieronymus *et al.*'s claim that ‘not just the use of the HDRS17 as measure of effect, but also many other methodological problems marring antidepressant trials, can be expected to make SSRIs appear less effective than they actually are’. Keeping the discussion on relative merits of different scales aside, the high prevalence rates of depression (Lewer *et al.*, 2015) decades after introduction of these drugs and enormous increase in the prescription rates of antidepressants (NHS digital, 2016) indicate that antidepressants are not as effective as they were thought to be.

Concluding remarks

We do not agree with Hieronymus *et al.*'s conclusion that there were inaccuracies, misleading statements and bias in our responses (Hieronymus *et al.*, 2018a). However, we acknowledge that in our response to their earlier critique (Katakam *et al.*, 2018), we were not able to clarify clearly some of the issues raised by Hieronymus *et al.* (2018b). We have now clarified our responses

in detail in this present response. We do not agree with Hieronymus *et al.*'s claim that our analysis regarding treatment of hepatitis C is flawed, and we wonder why they judge our analysis as flawed? We also wonder how they can conclude that interest in Cochrane checklists and handbooks can never substitute for actual insight into the subject of study? In conclusion, after accepting Hieronymus *et al.*'s valid suggestions for amendments, our updated analyses confirm our previous findings and conclusions. The harmful effects of SSRIs seem to outweigh the minimal (or non-existing) beneficial effects that SSRIs might have. Absence of evidence for harmful effects in young adults is not a valid evidence for absence of harmful effects in this age segment considering that SSRIs seem to raise SAEs in children (Olsson *et al.*, 2006; Sharma *et al.*, 2016; Gøtzsche, 2017) as well as the elderly (Katakam *et al.*, 2018). Regulatory guidelines do not request statistical significance before they consider advise against interventions (European Medicines Agency, 2017).

Supplementary material. To view supplementary material for this article, please visit <https://doi.org/10.1017/neu.2019.24>.

Acknowledgements. The authors thank Copenhagen Trial Unit, Centre for Clinical Intervention Research, for providing financial support. We have not received funding from any other sources.

Conflicts of interest. None.

References

- Adamson SJ, Sellman JD, Foulds JA, Frampton CM, Deering D, Dunn A, Berks J, Nixon L and Cape G (2015) A randomized trial of combined citalopram and naltrexone for nonabstinent outpatients with co-occurring alcohol dependence and major depression. *Journal of Clinical Psychopharmacology* **35**, 143–149.
- American Psychiatric Association (1980) *Diagnostic and Statistical Manual of Mental Disorders-DSM III*. Washington, DC.
- American Psychiatric Association (1987) *Diagnostic and Statistical Manual of Mental Disorders-DSM III-R*. Washington, DC.
- American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders-DSM IV*. Washington, DC.
- Arroll B, Elley CR, Fishman T, Goodyear-Smith FA, Kenealy T, Blashki G, Kerse N and MacGillivray S (2009) Antidepressants versus placebo for depression in primary care. *Cochrane Database of Systematic Reviews* **2009**, CD007954.
- Bagul NB and Kirkham JJ (2012) The reporting of harms in randomized controlled trials of hypertension using the CONSORT criteria for harm reporting. *Clinical and Experimental Hypertension* **34**, 548–554.
- Ball WA, Snively DB, Hargreaves RJ, Szegedi A, Lines C and Reines SA (2014) Addition of an NK1 receptor antagonist to an SSRI did not enhance the antidepressant effects of SSRI monotherapy: results from a randomized clinical trial in patients with major depressive disorder. *Human Psychopharmacology* **29**, 568–577.
- Boesen K, Paludan-Müller AS and Munkholm K (2018) Network meta-analysis of antidepressants. *Lancet* **392**, 1011–1012.
- Center for Drug Evaluation and Research (2018) *Major Depressive Disorder: Developing Drugs for Treatment Guidance for Industry - DRAFT GUIDANCE*. U. S. Food and Drug Administration. Available at <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM611259.pdf> (accessed 8 November 2018).
- Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, Leucht S, Ruhe HG, Turner EH, Higgins JP and Egger M (2018) Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet* **391**, 1357–1366.
- Claghorn JL, Earl CQ, Walczak DD, Stoner KA, Wong LF, Kanter D and Houser VP (1996) Fluvoxamine maleate in the treatment of depression: a

- single-center, double-blind, placebo-controlled comparison with imipramine in outpatients. *Journal of Clinical Psychopharmacology* **16**, 113–120.
- ClinicalTrials.gov** (2011) *Study Results for NCT00654953*. Available at <https://clinicaltrials.gov/ct2/show/results/NCT00654953?sect=X30156> (accessed 09 September 2018).
- Committee for Medicinal Products for Human Use** (2013) *Guideline on Clinical Investigation of Medicinal Products in the Treatment of Depression*. European Medicines Agency. Available at https://www.ema.europa.eu/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-depression_en.pdf (accessed 8 November 2018).
- Eli Lilly** (2004a) *Clinical Study Summary: Study F1J-MC-HMAQa*. Available at https://assets.contentful.com/hadumfdtzsr/3FeEMX77nykUwsguMgae/4f385628fa4e88bf5ec62fca7aedc23d/Duloxetine-F1J-MC-HMAQ-_Study-Group-A_.pdf (accessed 09 September 2018).
- Eli Lilly** (2004b) *Clinical Study Summary: Study F1J-MC-HMAQb*. Available at https://assets.contentful.com/hadumfdtzsr/5FZ15aVRBKKsmKIoEsgsMO/b1b988aefb263aabdb413af7f3534a7/Duloxetine-F1J-MC-HMAQ-_Study-Group-B_.pdf (accessed 09 September 2018).
- Eli Lilly** (2004c) *Clinical Study Summary: Study F1J-MC-HMATa*. Available at https://assets.contentful.com/hadumfdtzsr/1D76MTikKkeStuOS4SgQc/43330f08ad190a430f5d0f2140cb4b3/Duloxetine-F1J-MC-HMAT-_Study-Group-A_.pdf (accessed 09 September 2018).
- Eli Lilly** (2004d) *Clinical Study Summary: Study F1J-MC-HMATb*. Available at https://assets.contentful.com/hadumfdtzsr/5CF1OfIxBu0iwq8MWwEGOY/00f0112ecaeb9e28c5ace5c2c6087f4/Duloxetine-F1J-MC-HMAT-_Study-Group-B_.pdf (accessed 09 September 2018).
- Eli Lilly** (2014e) *ClinicalTrials.gov– Study Results for NCT01808612*. Available at <https://clinicaltrials.gov/ct2/show/results/NCT01808612> (accessed 09 September 2018).
- EudraCT Clinical Trial Results – CL3-01574-237** (2016) *EudraCT Clinical Trial Results – CL3-01574-237*. Available at <https://www.clinicaltrialsregister.eu/ctr-search/trial/2012-005612-26/results> (accessed 09 September 2018).
- European Medicines Agency** (2017) *Guideline on Multiplicity Issues in Clinical Trials*. European Medicines Agency. Available at https://www.ema.europa.eu/en/documents/scientific-guideline/draft-guideline-multiplicity-issues-clinical-trials_en.pdf (accessed 12 March 2019).
- Eyding D, Lelgemann M, Grouven U, Härter M, Kromp M, Kaiser T, Kerekes MF, Gerken M and Wieseler B** (2010) Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials. *British Medical Journal* **341**, e4737.
- Feighner JP and Overo K** (1999) Multicenter, placebo-controlled, fixed-dose study of citalopram in moderate-to-severe depression. *The Journal of Clinical Psychiatry* **60**, 824–830.
- Feighner JP, Robins E, Guze SB, Woodruff RA, Winokur G and Munoz R** (1972) Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry* **26**, 57–63.
- Forest Laboratories Inc** (2001) Study no. SCT-MD-01. Unpublished Data, Study Report from U.S. Food and Drug Administration.
- Fournier JC, Derubeis RJ, Hollon SD, Dimidjian S, Amsterdam JD, Shelton RC and Fawcett J** (2010) Antidepressant drug effects and depression severity: a patient-level meta-analysis. *Journal of the American Medical Association* **303**, 47–53.
- Furukawa TA, Salanti G, Atkinson LZ, Leucht S, Ruhe HG, Turner EH, Chaimani A, Ogawa Y, Takeshima N, Hayasaka Y and Imai H** (2016) Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. *British Medical Journal* **6**, e010919.
- Gastpar M, Singer A and Zeller K** (2006) Comparative efficacy and safety of a once-daily dosage of hypericum extract STW3-VI and citalopram in patients with moderate depression: a double-blind, randomised, multicentre, placebo-controlled study. *Pharmacopsychiatry* **39**, 66–75.
- Gibbons RD, Hur K, Brown CH, Davis JM and Mann JJ** (2012) Benefits from antidepressants: synthesis of 6-week patient-level outcomes from double-blind placebo-controlled randomized trials of fluoxetine and venlafaxine. *Archives of General Psychiatry* **69**, 572–579.
- GlaxoSmithKline** (2005a) *Scientific Result Summary – 29060.115*. Available at <https://www.gsk-clinicalstudyregister.com/files2/2060.pdf> (accessed 09 September 2018).
- GlaxoSmithKline** (2005b) *Scientific Result Summary – 29060.251*. Available at <https://www.gsk-clinicalstudyregister.com/files2/1989.pdf> (accessed 09 September 2018).
- GlaxoSmithKline** (2005c) *Scientific Result Summary – 29060.279*. Available at <https://www.gsk-clinicalstudyregister.com/files2/2866.pdf> (accessed 09 September 2018).
- GlaxoSmithKline** (2005d) *Scientific Result Summary – 29060.785*. Available at <https://www.gsk-clinicalstudyregister.com/files2/2202.pdf> (accessed 27 October 2017).
- GlaxoSmithKline** (2005e) *Scientific Result Summary – 29060.810*. Available at <https://www.gsk-clinicalstudyregister.com/files2/2204.pdf> (accessed 27 October 2017).
- GlaxoSmithKline** (2005f) *Scientific Result Summary – PAR29060.442*. Available at <https://gsk-clinicalstudyregister.com/files2/2323.pdf> (accessed 09 September 2018).
- Goldstein DJ, Lu Y, Detke MJ, Wiltse C, Mallinckrodt C and Demitrack MA** (2004) Duloxetine in the treatment of depression: a double-blind placebo-controlled comparison with paroxetine. *Journal of Clinical Psychopharmacology* **24**, 389–399.
- Goldstein DJ, Mallinckrodt C, Lu Y and Demitrack MA** (2002) Duloxetine in the treatment of major depressive disorder: a double-blind clinical trial. *Journal of Clinical Psychopharmacology* **63**, 225–231.
- Gotzsche PC** (2017) Antidepressants increase the risk of suicide, violence and homicide at all ages. *British Medical Journal* **358**, j3697.
- Gotzsche PC** (2018) *Rewarding the Companies that Cheated the Most in Antidepressant Trials [Internet]*. Mad in America. Available at <https://www.madinamerica.com/2018/03/do-antidepressants-work-a-peoples-review-of-the-evidence> (accessed 28 September 2018).
- Haidich A-B, Birtsou C, Dardavessis T, Tirodimos I and Arvanitidou M** (2011) The quality of safety reporting in trials is still suboptimal: survey of major general medical journals. *Journal of Clinical Epidemiology* **64**, 124–135.
- Hieronimus F, Emilsson JF, Nilsson S and Eriksson E** (2016) Consistent superiority of selective serotonin reuptake inhibitors over placebo in reducing depressed mood in patients with major depression. *Molecular Psychiatry* **21**, 523–530.
- Hieronimus F, Lisinski A, Näslund J and Eriksson E** (2018a) Katakam and co-workers have not shown SSRIs to be harmful and ineffective and should stop claiming that they have. *Acta Neuropsychiatrica* **2018**, 1–9.
- Hieronimus F, Lisinski A, Näslund J and Eriksson E** (2018b) Multiple possible inaccuracies cast doubt on a recent report suggesting selective serotonin reuptake inhibitors to be toxic and ineffective. *Acta Neuropsychiatrica* **30**, 244–250.
- Higgins JP and Green S** (2011) *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*. The Cochrane Collaboration. Available at <http://www.cochrane-handbook.org> (accessed 09 September 2018).
- Hodkinson A, Kirkham JJ, Tudur-Smith C and Gamble C** (2013) Reporting of harms data in RCTs: a systematic review of empirical assessments against the CONSORT harms extension. *BMJ Open* **3**, e003436.
- Hróbjartsson A, Emanuelsson F, Skou Thomsen AS, Hilden J and Brorson S** (2014) Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. *International Journal of Epidemiology* **43**, 1272–1283.
- Hróbjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I, Ravaut P and Brorson S** (2012) Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. *British Medical Journal* **344**, e1119.
- Hróbjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I, Ravaut P and Brorson S** (2013) Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. *Canadian Medical Association Journal* **185**, E201–E211.
- Hughes S, Cohen D and Jaggi R** (2014) Differences in reporting serious adverse events in industry sponsored clinical trial registries and journal articles on antidepressant and antipsychotic drugs: a cross-sectional study. *BMJ Open* **4**, e005535.

- ICH-GCP (1996) International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. (ICH-GCP). *ICH Harmonized Tripartite Guideline for Good Clinical Practice E6(R1)*. ICH-GCP. Available at https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf (accessed 13 October 2017).
- Ioannidis JP (2008) Effectiveness of antidepressants: an evidence myth constructed from a thousand randomized trials? *Philosophy, Ethics, and Humanities in Medicine* 3, 14.
- Ioannidis JP (2009) Adverse events in randomized trials: neglected, restricted, distorted, and silenced. *Archives of Internal Medicine* 169, 1737–1739.
- Jakobsen JC, Katakam KK, Schou A, Hellmuth SG, Stallknecht SE, Leth-Møller K, Iversen M, Banke MB, Petersen IJ, Klingenberg SL and Krogh J (2017a) Selective serotonin reuptake inhibitors versus placebo in patients with major depressive disorder. A systematic review with meta-analysis and Trial Sequential Analysis. *BMC Psychiatry* 17, 58.
- Jakobsen JC, Lindschou J, Hellmuth S, Schou A, Krogh J and Gluud C (2013a) The effects of selective serotonin reuptake inhibitors versus no intervention, placebo, or 'active' placebo in patients with major depressive disorder. A systematic review of randomised clinical trials with meta-analyses and trial sequential analyses. *PROSPERO*, CRD42013004420. Available at http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42013004420
- Jakobsen JC, Naqash JS and Gluud C (2017b) *Forskerzonen [Internet]. Copenhagen: Videnskab.dk; 2017. Forskere: Lægemiddelstyrelsens kritik af 'lykkepille'-studie holder ikke [Researchers: The critique by the Danish Medical Agency of study on the 'happiness pill' does not hold for scrutiny]; 26 June 2017 [cited 2019 Mar 15]; Available at <https://videnskab.dk/krop-sundhed/forskere-laegemiddelstyrelsens-kritik-af-lykkepille-studie-holder-ikke>.*
- Jakobsen JC, Simonsen E, Rasmussen KB and Gluud C (2013b) Is the total score of the Hamilton Depression Rating Scale associated with suicide attempts or suicides? *British Journal of Medical Research* 3, 140–152.
- Jindal RD, Friedman ES, Berman SR, Fasiczka AL, Howland RH and Thase ME (2003) Effects of sertraline on sleep architecture in patients with depression. *Journal of Clinical Psychopharmacology* 23(6), 540–548.
- Katakam KK, Sethi NJ, Jakobsen JC and Gluud C (2018) Great boast, small roast on effects of selective serotonin reuptake inhibitors: response to a critique of our systematic review. *Acta Neuropsychiatrica* 30(5), 251–265.
- Keller M, Montgomery S, Ball W, Morrison M, Snively D, Liu G, Hargreaves R, Hietala J, Lines C, Beebe K and Reines S (2006) Lack of efficacy of the substance p (neurokinin1 receptor) antagonist aprepitant in the treatment of major depressive disorder. *Biological Psychiatry* 59, 216–223.
- Khan A, Leventhal RM, Khan SR and Brown WA (2002) Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database. *Journal of Clinical Psychopharmacology* 22, 40–45.
- Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ and Johnson BT (2008) Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Medicine* 5(2), e45.
- Leucht S, Fennema H, Engel R, Kaspers-Janssen M, Lepping P and Szegedi A (2013) What does the HAMD mean? *Journal of Affective Disorders* 148, 243–248.
- Lewer D, O'reilly C, Mojtabai R and Evans-Lacko S (2015) Antidepressant use in 27 European countries: associations with sociodemographic, cultural and economic factors. *The British Journal of Psychiatry* 207, 221–226.
- Liu KS, Snively DB, Ball WA, Lines CR, Reines SA and Potter WZ (2008) Is bigger better for depression trials? *Journal of Psychiatric Research* 42, 622–630.
- Lundbeck (2005) *Clinical Trial Report Summary - Study 99024*. Available at http://www.lundbeck.com/upload/trials/files/pdf/completed/99024_CTRS_final_30Dec2005.pdf.
- Lundh A, Lexchin J, Mintzes B, Schroll JB and Bero L (2017) Industry sponsorship and research outcome. *Cochrane Database of Systematic Reviews* 2, MR000033.
- Mancino MJ, Mcgaugh J, Chopra MP, Guise JB, Cargile C, Williams DK, Thostenson J, Kosten TR, Sanders N and Oliveto A (2014) Clinical efficacy of sertraline alone and augmented with gabapentin in recently abstinent cocaine-dependent patients with depressive symptoms. *Journal of Clinical Psychopharmacology* 34, 234–239.
- Massana J (1998) Reboxetine versus fluoxetine: an overview of efficacy and tolerability. *The Journal of Clinical Psychiatry* 59, 8–10.
- Mathews M, Gommoll C, Chen D, Nunez R and Khan A (2015) Efficacy and safety of vilazodone 20 and 40 mg in major depressive disorder: a randomized, double-blind, placebo-controlled trial. *International Clinical Psychopharmacology* 30, 67.
- Moncrieff J (2018) *Challenging the New Hype About Antidepressants*. [Internet]. Clinical Psychiatry. Available at <https://joannamoncrieff.com/2018/02/24/challenging-the-new-hype-about-antidepressants> (accessed 28 September 2018).
- Moncrieff J and Kirsch I (2015) Empirically derived criteria cast doubt on the clinical significance of antidepressant-placebo differences. *Contemporary Clinical Trials* 43, 60–62.
- NHS Digital. digital.nhs.uk (2016) *Antidepressants were the Area with Largest Increase in Prescription Items in 2016*. Available at <https://digital.nhs.uk/news-and-events/news-archive/2017-news-archive/antidepressants-were-the-area-with-largest-increase-in-prescription-items-in-2016> (accessed 8 November 2018).
- Novartis (2009) *Clinical Trial Report for CAGO178A2303*. Available at <http://www.novctrd.com/ctrdWebApp/clinicaltrialrepository/displayFile.do?trialResult=2659> (accessed 09 September 2018).
- TV 2 (Denmark) (2017) *Nyt studie: 250.000 danskere tager pillerne – men de har ingen 873 effekt og mange alvorlige bivirkninger [250, 000 Danes takes pills – but they 874 show no beneficial effects and are connected with many serious adverse 875 events], [television broadcast]. Copenhagen; TV 2*. Available at <http://nyheder.tv2.dk/samfund/2017-02-09-nyt-studie-250000-danskere-tager-pillerne-men-de-har-ingen-effekt-og-mange>.
- Olsson M, Marcus SC and Shaffer D (2006) Antidepressant drug therapy and suicide in severely depressed children and adults: A case-control study. *Archives of General Psychiatry* 63(8), 865–872.
- Péron J, Maillet D, Gan HK, Chen EX and You B (2013) Adherence to CONSORT adverse event reporting guidelines in randomized clinical trials evaluating systemic cancer therapy: a systematic review. *Journal of Clinical Oncology* 31, 3957–3963.
- Pettinati HM, Oslin DW, Kampman KM, Dundon WD, Xie H, Gallis TL, Dackis CA and O'Brien CP (2010) A double-blind, placebo-controlled trial combining sertraline and naltrexone for treating co-occurring depression and alcohol dependence. *The American Journal of Psychiatry* 167, 668–675.
- Pfizer (2008a) *ClinicalTrials.gov – Study NCT00406952*. Available at <https://clinicaltrials.gov/ct2/show/NCT00406952?term=NCT00406952&rank=1> (accessed 09 September 2018).
- Pfizer (2008b) *ClinicalTrials.gov – Study NCT00636246*. Available at <https://clinicaltrials.gov/ct2/show/NCT00636246?term=NCT00636246&rank=1> (accessed 09 September 2018).
- Pfizer (2008c) *PhRMA Web Synopsis for Protocol A0501075 (NCT00636246)*. Available at http://www.rxarchives.org/uploads/2/4/4/6/24466638/a0501075_online.pdf (accessed 09 September 2018).
- Pfizer (2008d) *PhRMA Web Synopsis for Protocol A9001337(NCT00406952)*. Available at http://www.rxarchives.org/uploads/2/4/4/6/24466638/a9001337_online.pdf (accessed 09 September 2018).
- Pharmacia & Upjohn (2001a) *Reboxetine (PNU-155950E) Versus Placebo and Fluoxetine in a Controlled, Randomized, Double-Blind, Multicenter Study of Treatment in Major Depressive Disorders: Final Report of the Study Protocol 97-CRBX-050*. Available at http://www.iqwig.de/download/Studie_050.pdf (accessed 09 September 2018).
- Pharmacia & Upjohn (2001b) *Reboxetine, Placebo, and Paroxetine Comparison in Patients with Major Depressive Disorder: A Phase III, Randomized, Double-Blind, Placebo- and Active Treatment-Controlled, Parallel-Group, 8-Week Study of Reboxetine, Given Orally Twice Daily to Adult Patients with Major Depressive Disorder; Final Report of The Study Protocol M/2020/0046*. Available at http://www.iqwig.de/download/Studie_046.pdf (accessed 09 September 2018).
- Pharmacia & Upjohn (2001c) *Reboxetine, Placebo, and Paroxetine Comparison in Patients with Major Depressive Disorder: A Phase III, Randomized, Double-Blind, Placebo- and Active treatment-Controlled, Parallel-Group, 8-Week Study of Reboxetine, Given Orally Twice Daily To Adult Patients with Major Depressive Disorder: Final Report of The Study Protocol M/2020/0047*.

- Available at http://www.iqwig.de/download/Studie_047.pdf (accessed 09 September 2018).
- Ravindran AV, Teehan MD, Bakish D, Yatham L, O'reilly R, Fernando ML, Manchanda R, Charbonneau Y and Butters J** (1995) The impact of sertraline, desipramine, and placebo on psychomotor functioning in depression. *Human Psychopharmacology* **10**, 273–281.
- Savović J, Jones HE, Altman DG, Harris RJ, Juni P, Pildal J, Als-Nielsen B, Balk EM, Gluud C, Gluud LL and Ioannidis JP** (2012) Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Annals of Internal Medicine* **157**, 429–438.
- Sharma T, Guski LS, Freund N and Göttsche PC** (2016) Suicidality and aggression during antidepressant treatment: systematic review and meta-analyses based on clinical study reports. *British Medical Journal* **352**, i65.
- Shukralla AA, Tudur-Smith C, Powell GA, Williamson PR and Marson AG** (2011) Reporting of adverse events in randomised controlled trials of anti-epileptic drugs using the CONSORT criteria for reporting harms. *Epilepsy Research* **97**, 20–29.
- Smith SM, Chang RD, Pereira A, Shah N, Gilron I and Katz NP** (2012) Adherence to CONSORT harms-reporting recommendations in publications of recent analgesic clinical trials: an ACTTION systematic review. *PAIN* **153**, 2415–2421.
- Sundhedsstyrelsen [The Danish National Board of Health]** (2007) *Referenceprogram for unipolar depression hos voksne [Reference program for unipolar depression in adults]*. Sundhedsstyrelsen. Available at <http://www.sst.dk/~media/6F9CE14B6FF245AABCD222575787FEB7.ashx> (accessed 8 November 2018).
- Sutton AJ, Duval S, Tweedie R, Abrams KR and Jones DR** (2000) Empirical assessment of effect of publication bias on meta-analyses. *British Medical Journal* **320**, 1574–1577.
- Sweeting MJ, Sutton AJ and Lambert PC** (2004) What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Statistics in Medicine* **23**, 1351–1375.
- Timimi S, Moncrieff J, Gotzsche P, Davies J, Kinderman P, Byng R, Montagu L and Read J** (2018) Network meta-analysis of antidepressants. *Lancet* **392**, 1011–1012.
- Turner EH, Matthews AM, Linardatos E, Tell RA and Rosenthal R** (2008) Selective publication of antidepressant trials and its influence on apparent efficacy. *The New England Journal of Medicine* **358**, 252–260.
- Undurraga J and Baldessarini RJ** (2012) Randomized, placebo-controlled trials of antidepressants for acute major depression: thirty-year meta-analytic review. *Neuropsychopharmacology* **37**, 851.
- Warren J** (2018) Network meta-analysis of antidepressants. *Lancet* **392**, 1010–1011.
- Wernicke JF, Dunlop SR, Dornseif BE, Bosomworth JC and Humbert M** (1988) Low-dose fluoxetine therapy for depression. *Psychopharmacology Bulletin* **24**, 183–188.
- Whitaker R** (2018) *Do Antidepressants Work? A People's Review of the Evidence*. [Internet]. Mad in America. Available at <https://www.madinamerica.com/2018/03/do-antidepressants-work-a-peoples-review-of-the-evidence> (accessed 28 September 2018).
- World Health Organization** (1993) *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. Geneva, Switzerland: WHO.