Commentary

Multiple possible inaccuracies cast doubt on a recent report suggesting selective serotonin reuptake inhibitors to be toxic and ineffective

Hieronymus F, Lisinski A, Näslund J, Eriksson E. Multiple possible inaccuracies cast doubt on a recent report suggesting selective serotonin reuptake inhibitors to be toxic and ineffective.

According to a systematic review on the use of selective serotonin reuptake inhibitors (SSRIs) in adult depression that was recently published in BMC Psychiatry, the results of which have been widely disseminated in lay media, these drugs increase the risk for serious adverse events (SAEs) while exerting poor antidepressant efficacy. A cursory analysis, however, suggests the analysis of SAEs conducted by the authors to be marred by both methodological inaccuracies and blatant errors. After having corrected for these apparent mistakes, we conducted a sensitivity analysis in which we also accounted for a possible moderating effect of age; while this suggests SSRIs to be safe drugs in the non-elderly, they do confirm what is already known, that is, that they may enhance the risk for SAEs in the old. Given the loose definition of SAE, including also innocuous phenomena, the possible clinical significance of the latter observation, however, remains unclear until the nature and actual impact of the SAEs in question have been clarified. Moreover, with respect to efficacy, we find the paper in BMC Psychiatry misleading: first, the authors seem unaware of the well-established shortcomings associated with the conventional efficacy parameter on which their analysis is based, second, they have included suboptimal SSRI doses and third, they have missed some pivotal trials. Unless there are explanations for the many peculiarities in this paper that have escaped us, and which may be satisfactorily clarified by the authors, it seems important that the conclusions presented in this paper be publicly rectified.

On the basis of a systematic review and meta-analysis, Jakobsen et al. (1) have suggested that the toxicity of the selective serotonin reuptake inhibitors (SSRIs) outweighs an allegedly small antidepressant effect. If true, their conclusion should justify a drastic reconsideration of the current use of these drugs; if untrue, it may harm patients by making them refrain from accepting a potentially life-saving treatment. Prompted by the importance of the issue at stake, we have deemed it justified to cursorily scrutinise the analyses underlying the claims by Jakobsen et al. with particular focus on the one observation in their paper that may be regarded as truly novel, that is, that SSRIs are more likely to cause serious adverse events (SAEs), not only in the elderly, which is already well established, but regardless of age. As will be discussed below, we conclude that the calculations supporting this statement are marred by a number of peculiarities that, unless satisfactorily explained, are bound to cast serious doubts on the results.

It should be noted that the picture presented by Jakobsen et al. with respect to the issue of SSRI-induced harm is a mixed one. On the one hand, it is encouraging that their comprehensive investigation...
Hieronymus et al.

revealed no evidence for the oft-repeated claim that these drugs enhance the risk of suicide (2). This observation, notwithstanding the well-established methodological shortcomings marring the reporting of suicide-related adverse events in clinical trials, is in fact of such interest that it is most unfortunate that the authors refrain from mentioning it in the abstract and commenting on it in the discussion part of their paper. On the other hand, they do find that SAEs (without further specification) are more prevalent in SSRI-treated subjects than in those given placebo, such an event being observed in 31 out of 1000 patients treated with an SSRI and in 22 out of 1000 treated with placebo. Below are listed some of the reasons why we believe that caution is required when interpreting this result:

1. Jakobsen et al. give the reader the impression that their adverse event analyses are based on data from those patients that completed the full duration of each trial only. This would have been a questionable strategy since SAEs can obviously appear also in subjects discontinuing treatment prematurely; fortunately, the authors, however, do include SAEs occurring in drop-outs in their analyses. However, when calculating the risk for SAEs, they (often but not always) relate the number of SAEs among all patients starting treatment to the much smaller population of patients completing each trial. Needless to say, this manoeuvre will (i) inflate the apparent absolute risk of experiencing an SAE for both active treatment and placebo, and (ii) bias the results against whichever treatment displays the highest attrition rate.

2. The authors state that their analyses are based on the number of subjects experiencing at least one SAE. Whenever this information has been missing, they, however, seem to have assumed that the total number of reported SAEs corresponds to the number of subjects afflicted by at least one SAE (3), which will lead to an overestimation of the number of patients affected by SAEs. On the other hand, when the only SAE information accessible from the source used by the authors has concerned cases of deaths, rather than number of participants experiencing at least one SAE, it seems as if number of deaths has instead been used as the number of subjects experiencing SAEs, which may have led to an underestimation of the absolute risk of experiencing an SAE. For example, in table 2 of the Jakobsen report, the number of SAEs in the Kasper (4) study is claimed to be 2. This information seems to be taken from the published paper, where only deaths (n = 2) are presented, rather than from the Lundbeck study report (5), where 28 subjects are reported to have experienced at least one SAE.

3. Several studies (6–16) for which SAE data are readily available in the repositories used by the authors (17), and in some cases also in the published articles (6), are missing. Some of these trials (8–11) were zero-event studies, that is with no SAEs reported in any treatment group. Although excluding such studies from the calculation of odds ratios (ORs) may be reasonable (18,19), excluding them from the calculations of crude absolute risks will inflate the risk estimates.

4. One of the trials included by the authors, Pettinati (20), reports an SAE frequency of 37.5% in the SSRI monotherapy arm and of 28.2% in the placebo arm, hence constituting an extreme outlier with respect to SAE prevalence. The reason for this is obvious: this was a study in patients with depression combined with alcohol dependence, the most frequent SAE being ‘requiring inpatient detoxification and/or rehabilitation’. Needless to say, including this trial in a study of SSRI-induced SAEs will not increase our understanding of the tolerability of these drugs. Even more problematic is the handling of another outlier, the trial by Ravindran et al. (21), where the authors state that the frequency of SAEs is 16% in the SSRI arm and 15.4% in the placebo arm. A further analysis of possible reasons for these unexpected results would, however, have shown the authors that these numbers in fact do not refer to SAEs but to severe adverse events (i.e., any adverse event, including non-serious ones, that has been rated as severe in terms of intensity).

5. The take-home message of this paper, that is, that SSRIs should preferably be avoided because of their toxicity, is reinforced by the wording in the Method section and in the pre-published protocol, which suggests that the authors have made their own re-classification of adverse events, only regarding those that were ‘life threatening, resulted in death, disability, or significant loss of function, or caused hospital admission or prolonged hospitalisation’ as SAEs. Several of the adverse events reported in table 2 are, however, tabulated only as ‘unspecified SAE’ (n = 85) or ‘abnormal laboratory value’ (n = 17), suggesting that such a re-classification has in fact not been undertaken. Although the decision to refrain from reclassifying the SAEs may be justified by the inconsistent and limited reporting of harms in
journal articles (22,23), we find the current wording of the Method section on this matter misleading. Moreover, after having taken part of the sinister view provided by the first author when presenting his results in lay media, according to which dying, committing suicide and serious harm is what one may expect from the SSRIs (24–26), it is comforting to note that many of the events classified as SAEs, including nausea, missed abortion, unexpected pregnancy and appendicitis, are considerably less alarming (and/or of unlikely relationship to the given treatment). Conveying to patients and doctors the message that SSRIs cause enhanced risk for suicide appears particularly unjustified since the analyses of Jakobsen et al. provided no evidence whatsoever in support of such an association ($p=0.59$ for suicides; $p=0.31$ for suicide attempts).

6. A number of minor but significant errors seem to have found their way into the report: (i) The study by Ball et al. (27) did not include a placebo arm. The group reported as having been given placebo in table 2 in fact received a neurokinin-1 antagonist. (ii) A case of death in the trial by Nyth et al. (28) occurred during the single-blind lead-in but was regarded as an SAE during treatment. (iii) Female-specific SAEs in study GSK/810, which were more common in placebo-treated patients, seem to have escaped the attention of the authors, probably since they were presented in a separate table in the study report (29). (iv) For study NCT01473381 (30,31), which regards citalopram rather than (as stated in the Jakobsen paper) escitalopram, the reported OR appears incorrect (as judged by the numbers provided in table 2 and in the study report) (30). (v) The paroxetine 25-mg group of trial GSK/785 is reported as having four SAEs, whereas the correct number seems to be three (32). (vi) Study SCT-MD-01 is reported as having one citalopram arm, one escitalopram arm and one placebo arm but was in fact a four arm study including two escitalopram arms (33). (vii) Two comparisons from the three-arm Kasper (4) study are included in figure 11 but only one in table 2. Moreover, we are unable to replicate the ORs given in figure 11 for this trial regardless of whether we use the data from the paper (4) or from the study report (5). (viii) Two studies presented in figure 11, ‘99001, 2005’ and ‘Loo et al. (37)’, are, without obvious reason, missing in table 2. (ix) There is an unexplained mismatch with respect to the information regarding the total number of SAEs and the total number of patients provided in the text and in table 2, respectively (text: SSRI: 239/8242; placebo: 106/4956; table 2: SSRI: 211/7842; placebo: 94/4695). Including the two studies missing from table 2 does not resolve this discrepancy.

7. Contrary to the recommendations in the Cochrane handbook (34), which the authors otherwise seem most obedient to, they have, in most but not all (35) cases, chosen to report the different treatment arms from multi-arm studies separately in the SAE analysis. This practice is, however, unfortunate as it necessitates that the placebo group be subdivided so that participants are not counted twice, which in turn might bias the results as all included comparisons will not be independent (34). Furthermore, in cases where only one event occurred in the placebo group, this approach entails a risk of bias also by unnecessarily introducing continuity corrections in the meta-analysis.

8. Regarding continuity corrections, the confidence intervals and ORs provided in figure 11 appear excessive, as exemplified by Feighner (35), for which a 95% confidence interval ranging from $8.8e^{-14}$ to $5.7e^{18}$ and an OR of 704.14 are being reported. We would suggest these values to be the result of an inaccurate use of a continuity correction method, reciprocal zero-cell correction. What is described in the paper by Sweeting (18), to which Jakobsen et al. refer in this context, is the use of a continuity correction factor that is proportional to the ratio of patients in each treatment arm and which sums up to a constant, those of 0.01 or 1 being explored in the paper. According to Sweeting, their simulation studies suggest that, for unbalanced designs, a proportional constant that sums to 1 performs better than adding 0.5 to all cells (which is the conventional method); consequently Sweeting recommends that approach (among others). Although Sweeting’s method results in a continuity correction that is proportional to the imbalance in randomisation between arms, Jakobsen et al. seem to have taken the more literal approach of adding the reciprocal of the opposite treatment arm only to the cell with zero events. This strategy, however, leads to standard errors that are positively correlated to the size of the opposite treatment arm, implying that a trial with zero events in one arm is less informative the larger the size of its opposite arm, and also that increasing the size of the opposite arm directly decreases the risk of an event in the corrected arm. A related issue in need of clarification is that the authors state that ‘Review Manager version 5.3 was used for all meta-analyses’, whereas the graph of the SAE
Hieronymus et al.

analysis (figure 11) does not appear to be a product of this software. Moreover, the use of reciprocal zero-cell correction seems not to be mentioned in the pre-published protocol justifying the question if this be the result of a post-hoc adjustment of the intended SAE analysis, the pre-specified analysis remaining unreported?

Reading the referee comments regarding the Jakobsen paper, which BMC Psychiatry has made public, reveals that one of the reviewers suggested that the significant difference between SSRI's and placebo with respect to SAEs may be driven by an effect in the elderly. Invoking a lack of statistical heterogeneity and a paucity of heavily weighted trials in the meta-analysis, the authors, however, firmly rejected this possibility.

As there may well be clinically relevant differences between subgroups even in the absence of a verifiable overall statistical heterogeneity (36), and since previous studies suggest that there is indeed an increased propensity for certain possibly SSRI-related SAEs (e.g. osteoporosis and hyponatremia) in elderly patients, we, however, deemed the possibility raised by the reviewer deserving additional attention. Addressing this issue in a preliminary and cursory manner, and unfortunately without having access to all the data due to the study reports from Forest having been taken offline, we used RevMan 5.3 (Maentel-Haenszel random) to conduct an analysis aimed at assessing the possible impact of age while also addressing some of the methodological issues discussed above.

Basing our analyses on the data provided in table 2 in the paper by Jakobsen et al., we (i) added the two apparently erroneously non-tabulated studies (see above) (37,38), (ii) excluded studies that we believe should not have been included for reasons mentioned above (20,21,27), (iii) included missing treatment arms (see above), (iv) corrected all identified errors regarding the number of SAEs, (v) consistently used the intention to treat population for the patients at risk statistics and (vi) refrained from subdividing the placebo groups in multi-arm studies; for an annotated list of the changes we made to the data set, see Supplementary File 1. Although the overall result of this analysis reveals no significant difference between SSRI and placebo with respect to SAEs (OR 1.21, 0.94–1.56, p = 0.13; $I^2 = 0\%$), the test for subgroup differences was significant ($p = 0.049; I^2 = 74.1\%$), an increased risk for SAEs being observed in SSRI-treated subjects in studies regarding older subjects (OR 1.86, 1.13–3.06, $p = 0.01; I^2 = 0\%$) but no corresponding association found in the non-elderly studies (OR 1.04, 0.78–1.40, $p = 0.77; I^2 = 0\%$).

Although the outcome of this alternative analysis must be interpreted with caution as we had not access to all relevant data and made no attempt to include known missing data, and also as data extraction was conducted by merely one author (F.H.), it casts considerable doubt on the conclusion by Jakobsen et al. that SSRIs enhance the risk for SAEs regardless of age. Rather, it provides support for the conventional wisdom that SSRIs are generally safe drugs, with a very low propensity for SAEs in a non-elderly population, but that they, like many other relatively safe drugs, may cause SAEs in the elderly, as has since long been noted in the labelling for these drugs (39,40).

Although clarifying which specific SAEs that may be provoked by the SSRIs in the old, and their actual degree of seriousness, and to what extent there are specific risk populations for which these drugs perhaps should be avoided, may be warranted, propagating the inaccurate information that SSRIs are generally very harmful drugs that should be avoided is not.

Having analysed both adverse events and efficacy within the same project, Jakobsen et al. reached the somewhat arbitrary and subjective conclusion that ‘the potential small beneficial effects [of SSRIs] seem to be outweighed by harmful effects’. In this letter, we have focussed on methodological problems related to how the authors have analysed the alleged association between SSRIs and SAEs but refrained from discussing the issue of efficacy since the conclusions drawn in the Jakobsen paper in this regard are neither novel nor controversial: that SSRIs do outperform placebo, but that the difference between groups appears small when the effect is assessed using the sum score of the Hamilton Depression Rating Scale (HDRS), is hence in line with several previous meta-analyses. We, however, agree with Østergaard (41), who in a previous letter to this journal pointed out that the well-documented and well-argued shortcomings of the total rating of all 17 HDRS items as a measure of the antidepressant effect of SSRIs, disclosed, for example, by Per Bech et al. (for refs see 41), renders studies based on this parameter, including that of Jakobsen et al., grossly misleading.

Given that previous studies have firmly established that the mood-improving effect of SSRIs is not adequately captured when using the sum of the 17 different items of the HDRS as effect parameter, it is of note that Jakobsen et al. (as others have done before them), though basing their analysis on this insensitive measure, did obtain support for the superiority of SSRIs over placebo with respect to the risk of not achieving remission (relative risk (RR) 0.88, 0.84–0.91). Moreover, they reluctantly conclude that SSRIs ‘seem to significantly decrease the risk of no response’ (RR 0.83, 0.80–0.87), the $p$-value for this – seemingly – significant difference being reported as $p = 0.00001$ (but, as judged by the confidence limits, in fact probably being even lower). Appearing more loyal to their anti-SSRI beliefs than to their own results, the authors,
however, do not note these indications of SSRIs actually exerting an antidepressant effect with satisfaction or dispassionate rigour; instead, they blame themselves for not having conducted these analyses, regarding this as a ‘clear limitation’ of their report, and naming no less than four different post-hoc arguments for why it would have been preferable to keep themselves and their readers unaware of these differences. We suggest that the authors are overly self-critical in this regard: including assessments of remission and response does not constitute a limitation of their work but adds to provide a nuanced picture.

Finally, not only with respect to the SAE analyses, but also with respect to the assessment of efficacy, there seem to be some possible inaccuracies in the Jakobsen paper. The authors thus seem to have missed some of the pivotal positive SSRI trials (42,43), or included them in sensitivity analyses only (10–12,35,44–50), in spite of the fact that the information required to include them in the primary analyses should have been available to them, if not in the published papers, so in the Food and Drug Administration (FDA) reviews (51,52) and/or in the clinical trial summaries (10–12,45–50). Moreover, in the primary analysis of efficacy, the extremely low variance attributed to the study by Fabre (53) is probably an artefact: what are presented as standard deviations in the published paper are in fact, as judged by the trial report and the FDA review, standard errors rather than standard deviations (51,54). Finally, like previous meta-analyses attempting to assess the efficacy of the SSRIs, that of Jakobsen et al. is marred by the inclusion of treatment arms having received suboptimal doses of the tested SSRI (55).

To summarise, when scrutinising the paper by Jakobsen et al., we have identified a considerable number of problems and possible inaccuracies. It should be emphasised that these do not seem systematic or intentional; yet they do appear to have had a significant impact on the results and conclusions. Thus, unless there are explanations for these peculiarities that have escaped us when conducting this cursory review, and which may be satisfactorily clarified by the authors, it seems important that the information presented in the paper, as well as its conclusions, be rectified. With respect to the issue of toxicity, our own preliminary re-analysis of the relevant data suggests that SSRIs, as already known, may be associated with an increased risk for SAEs in the elderly, but that they appear quite safe in non-elderly subjects. With respect to efficacy, we believe that the pre-existing literature has convincingly shown these drugs to be clearly effective for the treatment of depression (41,55–57), and that the report by Jakobsen et al. supports rather than challenges this conclusion. Using their paper to propagate the image of the SSRIs as generally toxic and ineffective is hence misleading.

Possible inaccuracies in recent SSRI meta-analysis

Acknowledgements

None.

Conflicts of Interest

F.H. has received speaker’s fees from Servier. E.E. has previously been on advisory boards and/or received speaker’s honoraria and/or research grants from Eli Lilly, GlaxoSmithKline, Servier and Lundbeck. The other authors report no conflicts of interest.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/neu.2017.23

References

Hieronymus et al.

34. HEGINS JPT, GREEN S. Cochrane handbook for systematic reviews of interventions version 5.0.0, 16.5.4, 2011. Available at http://handbook-5-1.cochrane.org/chapter_16/16_5_4_how_to_include_multiple_groups_from_one_study.htm.
Possible inaccuracies in recent SSRI meta-analysis


