PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Incident User Cohort Study of Risk for Gastrointestinal Bleed and Stroke in Individuals with Major Depressive Disorder Treated with Antidepressants
AUTHORS	Victor Castro, Patience J. Gallagher, Shawn N. Murphy, Vivian S. Gainer, Maurizio Fava, Jeffrey Weilburg, Susanne E. Churchill, Isaac Kohane, Dan V. Iosifescu, Jordan W. Smoller and Roy H. Perlis

VERSION 1 - REVIEW

REVIEWER	Mark S. Bauer, MD Professor of Psychiatry, Harvard Medical School VA Boston Healthcare System USA Conflict of Interest: Though I am nominally in the same department as the authors, our "department" is actually comprised of more than a half-dozen different departments that are linked only by an administrative committee of chairs. While I know some of the authors professionally, I do not collaborate with any of them and had no knowledge of this study prior to receiving the request to review. I do not feel compromised in my assessment of this manuscript.
REVIEW RETURNED	16/12/2011

THE STUDY	Modest levels of concern as detailed in the critique for the authors.
GENERAL COMMENTS	The authors present a cohort study of the risk of gastrointestinal bleeding, stroke, or myocardial infarction in patients diagnosed with major depression treated with newer antidepressants with higher vs. lower affinity for the serotonin transporter. They find a modest but statistically significant association with the first two outcomes but not the third, and conclude that these results add to the growing literature indicating that some newer antidepressants with high affinity for the serotonin transporter are associated with increased risk of gastrointestinal bleed or stroke. What is new in this paper is the characterization of antidepressants as having higher vs. lower affinity, in order to investigate the role of serotonin transporter affinity within the group of antidepressants that have serotonin reuptake inhibiting characteristics. One assumes that a genotyping study is soon to follow. The paper is well written, and the results are presented well and discussed succinctly but cogently. There are several methodologic aspects of the paper that require clarification or revision before the results can be fully interpreted. 1. The incident cohort methodology is used with increasing frequency in such surveillance studies; one thinks also of the lithium-

	suicide study of Goodwin and colleagues several years ago. What is not clear from this report, though, is whether only one or multiple exposures in a single patient were included as observations. For instance, if a subject had 3 months of a high affinity drug and then experienced a hiatus, and subsequently received 6 months of another high affinity drug, were both epochs of treatment counted? Only the first? Other? What is the person-month count that this subject contributes? 2. It is not clear whether medication exposure derived the EMR was gleaned from actual prescription records (and if so, from clinician orders or actual fills), from progress note text, or from some other clinician-derived entry in the EMR. If it is not clear that the medication has actually been dispensed, this is a potential source of bias (likely deflating association, as the actual exposure of those listed as exposed would be less). 3. On a related matter, the authors investigate aspirin usage as a potential covariate. Since this is an over-the-counter medication, for which a prescription would not typically be written, how were these data captured, and what is the likely reliability? 4. More substantively, it is not clear why the authors did not investigate other NSAIDs as well as aspirin, which have been demonstrated in some studies to increase risk in the context of serotonin reuptake inhibitor antidepressants. 5. Of some methodologic concern is the identification of individuals with major depression by virtue of having a single diagnosis in EMR billing data. Most administrative database studies have utilized the criteria of one inpatient hospitalization or two outpatient visits with a given diagnosis as meeting criteria for the corresponding disorder. Using this higher threshold for diagnosis may have reduced the sample size unduly, and therefore using this lower bar would be a reasonable trade-off. However, if the investigators are not able to utilize the more standard definition, it is really not justifiable to list the use of only those with a major depression diagnosis as a strength—since it's such a low threshold for diagnosis that it is very likely that many of their subjects do not have what researchers, or even most clinicians, would consider true major depression. 6. Related to these concerns, the logic of excluding those with lower levels of depression in order to reduce the potential confound by indication is not tenable, since there has been an extensive literature linking major depression to increased cardiovascular risk. It is as if the authors are saying that the more severe the pathology the less the confounding by indication. It is best to drop this as a putative strength and, optimally, to use a higher threshold for diagnosing major depression if this is the group they wish the readers to draw conclusions regarding. 7. Further, was there a chronological relationship between the receipt of diagnosis and treatment exposure? For instance, could a subject have had a major depression diagnosis in 2003 and have their treatment exposure in 2007? Or did the diagnosis have to be gleaned from an interval of treatment exposure? Overall, though, with the clarifications suggested above and addressing the limitations either by additional data analyses or by presenting their conclusions with appropriate limitations, this paper

	can incrementally add to the literature.

REVIEWER	Rudolf Uher Clinical Lecturer Institute of Psychiatry, King's College London UK
REVIEW RETURNED	08/01/2012

THE STUDY	Additional reference for inclusion: Coupland C, Dhiman P, Morriss R, Arthur A, Barton G, Hippisley-Cox J. Antidepressant use and risk of adverse outcomes in older people: population based cohort study. BMJ 2011; 343:d4551.
RESULTS & CONCLUSIONS	The manuscript would be more informative if (a) the actual KD for each antidepressant is given in Table 1 and (b) the rates of adverse events are also reported for each antidepressant separately (at least for antidepressants used by >1000 individuals). This is important if someone wants to compare rates across studies or carry out a meta-analysis.
GENERAL COMMENTS	The manuscript by Castro and colleagues reports higher risk of gastrointestinal bleeds and stroke in patients taking antidepressants with strong serotonin reuptake blocking properties. Neither association is novel, but the large sample size and tight design make this an important confirmatory study. The study is generally well executed. The association with both ischaemic and haemorrhagic stroke may look surprising, but it is consistent with literature and it is well discussed. The authors may want to consider the following issues in revision of the manuscript: (1) The reported results are predicated on the division between weak and strong serotonin-reuptake blockers. Antidepressants are separated into the two groups through a cutoff of 1 on the dissociation constant (KD). Some of the classifications are surprising – e.g. citalopram and escitalopram end up on the opposite sides of the divide. The manuscript would be more informative if (a) the actual KD for each antidepressant is given in Table 1 and (b) the rates of adverse events are also reported for each antidepressant separately (at least for antidepressants used by >1000 individuals). This is important if someone wants to compare rates across studies or carry out a meta-analysis. (2) A large recent study by Coupland et al (BMJ 2011; 343:d4551) is missing from the reference list even though it reports both stroke and gastrointestinal bleeds. A comparison of these two large studies on methods and results may be useful. (3) Since one of the aims is to quantify the effect size, risk ratios may be more suitable than odds ratios and would be more comparable to hazard ratios reported in other studies.

VERSION 1 – AUTHOR RESPONSE

Reviewer: Mark S. Bauer, MD

1. What is not clear from this report, though, is whether only one or multiple exposures in a single patient were included as observations. For instance, if a subject had 3 months of a high affinity drug

and then experienced a hiatus, and subsequently received 6 months of another high affinity drug, were both epochs of treatment counted? Only the first? Other? What is the person-month count that this subject contributes?

Each exposure incidence contributes to a patient's exposure period and is counted in the person-month calculation. In the example above, both epochs would be counted and the patient would have a 9-month exposure window. Bleeding outcomes of interest occurring within this exposure period will be associated with the drug. A sentence has been added to the methods section of the manuscript to clarify this point.

2. It is not clear whether medication exposure derived the EMR was gleaned from actual prescription records (and if so, from clinician orders or actual fills), from progress note text, or from some other clinician-derived entry in the EMR. If it is not clear that the medication has actually been dispensed, this is a potential source of bias (likely deflating association, as the actual exposure of those listed as exposed would be less).

We have clarified that medication exposure comes from three sources: 1) medications prescribed to the patient via e-prescribing in the EMR, 2) medications documented but not prescribed by the documenting clinician and 3) medications dispensed by the inpatient pharmacy. As the reviewer notes, unfilled prescriptions (as well as poor adherence) would bias us away from detecting association; we have addressed this as a limitation in the discussion, including poor adherence and unfilled prescriptions.

3. On a related matter, the authors investigate aspirin usage as a potential covariate. Since this is an over-the-counter medication, for which a prescription would not typically be written, how were these data captured, and what is the likely reliability?

See point 2, above; non-prescription medications used on a regular basis are documented in the EMR even if no prescription is issued. Aspirin dispensed from the inpatient pharmacy is also captured. We cannot exclude the possibility of undersampling, though we note that the positive control was indeed positive.

4. More substantively, it is not clear why the authors did not investigate other NSAIDs as well as aspirin, which have been demonstrated in some studies to increase risk in the context of serotonin reuptake inhibitor antidepressants.
We have clarified in the text that aspirin was included solely as a positive control; it was not our intention to investigate any possible contributor to bleeding, which we felt would be beyond the scope of this report.

5. Of some methodologic concern is the identification of individuals with major depression by virtue of having a single diagnosis in EMR billing data. Most administrative database studies have utilized the criteria of one inpatient hospitalization or two outpatient visits with a given diagnosis as meeting criteria for the corresponding disorder. Using this higher threshold for diagnosis may have reduced the sample size unduly, and therefore using this lower bar would be a reasonable trade-off. However, if the investigators are not able to utilize the more standard definition, it is really not justifiable to list the use of only those with a major depression diagnosis as a strength—since it's such a low threshold for diagnosis that it is very likely that many of their subjects do not have what researchers, or even most clinicians, would consider true major depression.

We chose the broadest possible cohort to maximize our power to detect association, the trade-off noted by the reviewer, but recognize that our definition was broader than is traditional. We have added a comment to this effect. We have run a secondary analysis of primary outcomes using the

narrower 2+ diagnostic code definition for consistency with prior reports, and added a brief comment to the results and added a comment and a table (S4) to the appendix

6. Related to these concerns, the logic of excluding those with lower levels of depression in order to reduce the potential confound by indication is not tenable, since there has been an extensive literature linking major depression to increased cardiovascular risk. It is as if the authors are saying that the more severe the pathology the less the confounding by indication. It is best to drop this as a putative strength and, optimally, to use a higher threshold for diagnosing major depression if this is the group they wish the readers to draw conclusions regarding.

See #5, above; we have clarified our logic, which was simply intended to point out a limitation of prior reports such as Smoller, Arch Int Med 2009.

7. Further, was there a chronological relationship between the receipt of diagnosis and treatment exposure? For instance, could a subject have had a major depression diagnosis in 2003 and have their treatment exposure in 2007? Or did the diagnosis have to be gleaned from an interval of treatment exposure?

A diagnosis of major depression was used as a means of selecting the study population. Exposures and outcomes may have occurred prior to the initial diagnosis of major depression. However, all outcomes must have occurred after anti-depressant treatment.

Reviewer: Rudolf Uher
Additional reference for inclusion: Coupland C, Dhiman P, Morriss R, Arthur A, Barton G, Hippisley-Cox J. Antidepressant use and risk of adverse outcomes in older people: population based cohort study. BMJ 2011; 343:d4551.

The reference has been included.

The manuscript would be more informative if (a) the actual KD for each antidepressant is given in Table 1 and (b) the rates of adverse events are also reported for each antidepressant separately (at least for antidepressants used by >1000 individuals). This is important if someone wants to compare rates across studies or carry out a meta-analysis.

KD values have been added to Table 1. In addition, we have included a table (S2) in the supplemental materials with rates of bleed adverse events reported for each anti-depressant.

The authors may want to consider the following issues in revision of the manuscript:
(1) The reported results are predicated on the division between weak and strong serotonin-reuptake blockers. Antidepressants are separated into the two groups through a cutoff of 1 on the dissociation constant (KD). Some of the classifications are surprising – e.g. citalopram and escitalopram end up on the opposite sides of the divide. The manuscript would be more informative if (a) the actual KD for each antidepressant is given in Table 1 and (b) the rates of adverse events are also reported for each antidepressant separately (at least for antidepressants used by >1000 individuals). This is important if someone wants to compare rates across studies or carry out a meta-analysis.

As above, KD values added to Table 1 and added Table S2 in the supplemental materials with rates of bleed adverse events reported for each anti-depressant

(2) A large recent study by Coupland et al (BMJ 2011; 343:d4551) is missing from the reference list even though it reports both stroke and gastrointestinal bleeds. A comparison of these two large studies on methods and results may be useful.

As above, the reference has been included.

(3) Since one of the aims is to quantify the effect size, risk ratios may be more suitable than odds ratios and would be more comparable to hazard ratios reported in other studies.

We have amended the manuscript to report relative risk to enable more direct comparison to prior large cohort studies.

VERSION 2 – REVIEW

REVIEWER	Mark S. Bauer, MD Professor of Psychiatry Harvard Medical School VA Boston Healthcare System
REVIEW RETURNED	06/02/2012

The reviewer completed the checklist but made no further comments.