Author's response to reviews

Title: One-year risk of psychiatric hospitalization and associated treatment costs in bipolar disorder treated with atypical antipsychotics: a retrospective claims database analysis

Authors:

   Edward Kim (edward.kim@bms.com)
   Min You (min.you@bms.com)
   Andrei Pikalov (andreipikalov@otsuka.com)
   Quynh-Van Tran (Quynh-Van.Tran@otsuka.com)
   Yonghua Jing (yonghua.jing@bms.com)

Version: 3 Date: 12 November 2010

Author's response to reviews: see over
Dear Dr. Jing,

Your revised manuscript has now been re-reviewed and the comments are accessible in PDF format from the links below. Please let us know if you have any problems opening the files.

Referee 1: http://www.biomedcentral.com/imedia/1629649837400464_comment.pdf
Referee 3: http://www.biomedcentral.com/imedia/5282065283996597_comment.pdf
Referee 2: http://www.biomedcentral.com/imedia/9455873941645396_comment.pdf

**************************************************************************
**************************************************************************
Associate Editor comments:

Please provide a further minor revision paying close attention to the comments of Reviewer 1. Please provide answers to the concerns of reviewer 2, including further succinct justification in the manuscript text and/or a clear statement defending their presentation. It will be helpful if this can be highlighted within the revision.

**************************************************************************
**************************************************************************
Editorial comments:

As you will see there have been some further criticisms of your manuscript. Please note that it is our policy to allow a maximum of two revisions on manuscripts under consideration and thus the next revision is the last on which we are willing to seek advice. We therefore urge you to make every effort to fully address the criticisms during this revision.

Please replace 'Objectives' with 'Background' within your Abstract.

We would be grateful if you could address the latest comments in a revised manuscript and provide a cover letter giving a point-by-point response to the concerns.

Please also highlight (with 'tracked changes'/coloured/underlines/highlighted text) all changes made when revising the manuscript to make it easier for the Editors to give you a prompt decision on your manuscript.
Please also ensure that your revised manuscript conforms to the journal style (http://www.biomedcentral.com/info/ifora/medicine_journals). It is important that your files are correctly formatted.

We look forward to receiving your revised manuscript by 9 August 2010. If you imagine that it will take longer to prepare please give us some estimate of when we can expect it.

You should upload your cover letter and revised manuscript through http://www.biomedcentral.com/manuscript/login/man.asp?txt_nav=man&txt_man_id=1396617724317848. You will find more detailed instructions at the base of this email.

Please don't hesitate to contact me if you have any problems or questions regarding your manuscript.

With best wishes,

Angelina Ilievksa MSc
on behalf of Dr. Greenshaw.

Tel: +44 (0) 20 3192 2013
e-mail: editorial@biomedcentral.com
Web: http://www.biomedcentral.com/

Additional material submitted by the reviewers
Response to reviewer 1

Reviewer’s report:
Major Compulsory Revisions

There remain a number of outstanding issues with this manuscript itself. There are still a few questions that, in the spirit of scholarly transparency, should be answered. In addition, a few assertions about their methods need correction.

Page 6: The authors assert, “As the dataset in this study was derived from an insurance claim database, ethical approval and informed patient consent were not required.” Could the authors clarify whether this statement is based on the authors’ opinion or that of an ethics review board? Although the use of administrative data may not seem to require consent, it does not automatically mean that the research project does not need ethical approval from a review board.

Response: Because the data conform to the Health Insurance Portability and Accountability Act of 1996 confidentiality requirements, the study did not require informed consent or institutional review board approval. We clarified this in the revised manuscript.

Page 7: Why were those who had a hospitalization 7 days following an index prescription excluded from the analyses?

Response: All patients were outpatients at the time of the index date, and were censored if they were hospitalized within 7 days of their index antipsychotic prescription in order to eliminate patients who were so unstable at the time that no treatment might be effective – this has now been made clear in the methods.

Page 7: The exclusion of people who discontinued antipsychotics introduces a potential bias in the analyses. It is a limitation that should be discussed. The results only hold for those who do not discontinue use. There seems to be a large proportion of people who discontinue use of the drug that the analyses find less costly. Perhaps, the authors could test whether there is a significant discontinuation among the drugs.

Response: This information was added to the limitation section. “Finally, the study only followed patients until their first psychiatric hospitalization and did not address outcomes following adding, switching, or discontinuing antipsychotics, which may be common in this population.

The duration of therapy on atypical antipsychotics was comparable across all treatment groups and fairly brief, with a median of 30 days across all treatments.
Page 8: The authors have misunderstood reference 15. Propensity scores do not control for unobserved confounding variables. Only instrumental variables do this, and that is why they were asked it about in the previous review. The statement that references source #15 is incorrect, and source #15 did not state this contention either. In addition, in their response, the authors indicate that they include the variables that they did in the logistic regression because they were limited by the dataset. This is a potential limitation that should be noted and discussed.

Response: The statement was corrected and the reference was changed. The limitation that we could only include the variables available in the dataset was noted in the limitation section.

Page 8: The authors include more information about the Charlson index in their letter. To inform readers, they should include the additional information in the manuscript.

Response: Additional information about the Charlson Index was included in the appendix of the revised manuscript.

Page 9: The authors indicate that they use a gamma regression. To inform readers, they should also describe how they: (1) converted the coefficients to dollars, (2) calculated standard errors for the dollars, (3) tested for heteroskedasticity, and (4) what the results of their test for heteroskedasticity were. As requested in the previous review, it is important to describe the fit of the model especially since they have already indicated that they were limited with regard to the types of variables available to them.

Response: First we calculated the mean for each of the numeric covariates, and gave equal share of the categorical covariates, and then calculated the log mean of the fitted gamma distribution based on these covariate values and the parameter estimates and exponentiated the log mean to get the cost in dollars. We got standard errors for dollars in the same way. We have clarified in the manuscript.

There is no fit statistic for gamma regression. We used propensity score matching method and included the covariates that would affect post-index costs to better the prediction model.

Page 13: They should indicate that their use of propensity score matching could have resulted in treatment selection bias based on unobserved variables. One of the drawbacks of propensity score matching is that it only balances the data that one has. They have already indicated that their data are limited. If there are unobserved factors that affect the choice of treatment, these will bias the results.
Response: This was acknowledged in the study limitations. The unmatched results were also displayed in the Appendix.

Page 14: The authors assert, “However, the analysis of such complex treatment patterns within claims data may be subject to high levels of unobservable confounding and difficult to interpret with respect to contribution of individual medications across complex regimens.” The authors should discuss the implications of this on the interpretation of their results.

Response: We have discussed the implications of this on the interpretation of the results.

Page 14: Is it possible that the patients in this dataset received treatment from other health or mental health programs that are not covered in this administrative dataset? If this is a possibility, that should also be noted as a limitation and discussed how this could affect the “total healthcare” costs that the authors report.

Response: All mental health claims are included in the data source, including any carve-out mental health services.

Table 1. The previous review asked about how the p-values were calculated. Was the matching accounted for or do the p-values ignore the fact that the patients have been matched?

Response: Although the patients are matched, there are still differences among the matched patients. The p-values were based on the gamma regression, controlling for co-variates.

Table 3. It would be helpful if the title indicated what the outcome is. In addition, it is “proportional” rather than “proportionate”. The authors might also consider including a graph of the curves from the results of the analysis presented in Table 3.

Response: The changes were made.
Response to reviewer 2

Reviewer’s report

The responses to the "Compulsory Issues" have not adequately addressed the issues. Furthermore, although the authors have made lengthy responses to the reviewers' comments, they have not indicated where any changes to the manuscript have been made.

Response: We responded to the reviewer’s queries, some of which appeared to request justification for our methods that we believe were already addressed in the manuscript. Requests to clarify specific details in the methods, results, and discussion were not identified. We are eager to respond to such queries on behalf of the readers of the journal.

1. Concerns were raised about "key statistical significance values" in Table 3. The authors indicate that the multiple comparisons within this kind of modeling are conducted independently and thus are not subject to this kind of concern (the reference is appreciated). It can be seen that comparing several treatments to one other Rx is like comparing them to zero - thus no issue. However, the standard of comparison here is also a treatment of interest, so it is not clear to me that the explanation provided by the authors applies. Thus, I have recommended a statistical review of the paper. In regard to the second point within this issue, i.e. the use of an alpha level of 0.05, that is up to the journal. However, the authors' assertion that "retrospective observational studies are inherently biased to the null hypothesis due to confounding by indication and the crudeness of measurements" may or may not apply to this study, and is thus not an adequate argument in this kind of study. It should be noted that if the authors' statement does apply, it is not a very good argument for its publishability.

Response: This comment does not request changes to the manuscript. We agree that a statistical review is indicated due to a difference of opinion on the methodology used. We are concerned about the tone of this comment, for which no edits to our manuscript can be coherently inferred.

Aripiprazole is the reference/control group in the Cox proportionate hazards models. We developed four Cox proportionate hazards models and each model was independent of the rest. In each model, only two drugs were included in the model (aripiprazole and one comparator). The samples in the four models are different and independent.
For comparisons to be valid, the composition of the reference group should resemble that of the comparator group as closely as possible. Our use of propensity score matching was an attempt to follow this principle. We defer to the journal’s threshold for statistical significance and believe that a statistical reviewer will confirm that several multivariate models with the same reference are not subject to a Bonferroni correction that would be required with multiple univariate comparisons.

2. The matter of six "differences" that were individually not significant, but all in the same direction, begs the question of whether they might have been statistically significant if taken together as a class (i.e. mood stabilizers) as it appears that they were in discussion. The authors may well be right in the end, but an overall analysis of mood stabilizers might have been useful. Furthermore, although attention to clinical contingencies is important, it is not reasonable to address a statistical/design issue with clinical conjecture (e.g. "...patients were more likely to be on concomitant mood stabilizers and therefore 'doubly' protected against hospitalization." This is a limitation, not a strength.

Response:
We believe that our use of individual mood stabilizers as variables for propensity score calculation addresses concerns about potential unequal use of mood stabilizers. The table below demonstrates that the analysis sample was well matched on mood stabilizer use. We also believe that all statistical modeling of health services data must be consistent with plausible clinical effects. This is a principle of clinical epidemiology and the source of this reviewer’s objections.

We believe that the reviewer has requested an additional analysis on whether use of mood stabilizers as a class differed among treatments. However, precisely how results of this analysis are to be incorporated into the manuscript has not been specified.

<table>
<thead>
<tr>
<th>Matching Scheme</th>
<th>Aripiprazole</th>
<th>Comparator</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matching aripiprazole to ziprasidone</td>
<td>245 (53.1%)</td>
<td>253 (54.9%)</td>
<td>0.5971</td>
</tr>
<tr>
<td>Matching olanzapine to aripiprazole</td>
<td>422 (57.3%)</td>
<td>410 (55.6%)</td>
<td>0.5284</td>
</tr>
<tr>
<td>Matching quetiapine to aripiprazole</td>
<td>457 (59.4%)</td>
<td>471 (61.2%)</td>
<td>0.4660</td>
</tr>
<tr>
<td>Matching risperidone to aripiprazole</td>
<td>456 (59.1%)</td>
<td>451 (58.5%)</td>
<td>0.7959</td>
</tr>
</tbody>
</table>

3. This issue is that the propensity-score matching exercise led to comparisons based on samples that were unequal and perhaps biased. The response did not
address this concern except to point to the limitations of the approach and to explain that "it would be counterintuitive to expect that all four comparisons would use the exact same patients". There is no disagreement about this analysis being difficult, but that is not an argument that deals with any potential biases.

Response:
On this matter we again defer to the statistical review that has been appropriately recommended by this reviewer. Propensity score matching is not intended to analyze the same samples but rather to minimize differences between the samples that are actually analyzed. This by definition requires exclusion of certain cases for which matching is not possible. Excluded cases will vary depending on the comparison in order to achieve this match. Hence there is no reason to expect that all comparisons will use the same aripiprazole patients. Moreover, sample variability is not the same as bias, so we are perplexed by this query.

Propensity score matching is limited by the smallest sample size. With all comparators except ziprasidone, aripiprazole was the smaller sample. We also performed multivariate analysis for the full unmatched samples as a sensitivity analysis. Those results are in the Appendix.