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 (andrei.pikalov@otsuka.com)
Quynh-Van Tran (Quynh-Van.Tran@otsuka.com)
Yonghua Jing (yonghua.jing@bms.com)

Version: 2 Date: 20 April 2010

Author's response to reviews: see over
20 April 2010

Dear Dr Alam,

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

Following receipt of your email with the reviewers comment for the above manuscript, we have revised the manuscript and now resubmit it for publication in BMC Psychiatry.

A point-by-point response to the reviewers’ comments is attached to this letter.

We hope that our manuscript is now suitable for publication and we look forward to hearing from you.

Yours sincerely,

Yonghua Jing, PhD

Bristol-Myers Squibb
777 Scudders Mill Road
Plainsboro
NJ 08536
USA
Email: yonghua.jing@bms.com
Tel: +1 609-897-2914
Fax: +1 609-897-6319
Response to Reviewer 1: Carolyn Dewa
Reviewer's report:
The study compared one-year risk of psychiatric hospitalization and treatment costs in commercially insured patients with bipolar disorder, treated with aripiprazole, ziprasidone, olanzapine, quetiapine or risperidone.

Major Compulsory Revisions
The main concern arises with the censoring of those who switched antipsychotics and those who discontinued use. Both groups may have represented people who may have more complex courses of treatment. This complex course could be associated with increased hospital use, but the censoring prevents the exploration of this possibility. Table 2 indicates that a large proportion of people who used antipsychotics discontinued use. It seems that the largest group to discontinue use were those who used aripiprazole. The authors seem to overlook this outcome, focusing instead on hospitalizations; however, the percentage of people who discontinued use was almost 10 times greater than the percentage with a psychiatric hospital admission.

Also, although adherence to medication is an important factor contributing to hospitalization, this was not adjusted for in the statistical analyses; in fact, subjects who were non-adherent were dropped from the analysis. In addition, the authors conclude that aripiprazole is associated with lower total healthcare costs. Yet, this conclusion must be qualified by the fact that the authors only looked at those who had positive costs. This raises the question of what results would have been observed if all subjects had been included in the analyses. Based on their selection criteria, the sample decreased from 198,919 to 7,169 people. The analyses are based on only 3.6% of the entire sample.

Response: The fact that patients were able to discontinue antipsychotic medication without subsequent psychiatric hospitalization during the observation period would indicate that the medication was no longer needed. APA practice guidelines for bipolar disorder recommend simplifying the medication regimen when the patient is stabilized. Therefore, unlike schizophrenia, where continued antipsychotic therapy is a core component of treatment, in bipolar disorder, discontinuation of antipsychotic (and maintenance on mood stabilizer) is a more standard treatment paradigm.

We dropped all patients who were non-adherent in order to develop a more robust estimate of the effects of medication on outcome. Including non-adherent patients would have further confounded the results.

Regarding sample size, the original 198,919 patients included all patients with at least one prescription for an antipsychotic. The exclusion of patients with non-positive costs is a common approach in that it typically reflects erroneous data, i.e. a patient has received a prescription yet has no evidence of having received billable medical services. This is not logical and therefore makes analysis of such a patient questionable.
Specific comments
Page 9, Paragraph 1: Why were the variables chosen for the logistic regression? Is there a rationale for them? There does not seem to be adjustment for the length of time a person had bipolar disorder.

**Response:** The rationale for the variables is that they are available and likely to contribute to treatment outcomes. There is regional variability in hospitalization patterns, and choice of mood stabilizer is an important potential confounder.

Bipolar disorder is a chronic condition whose duration may exceed the enrollment period of the patient. We have no way of verifying the duration of disease.

Page 9, Paragraph 1: How were the p-values calculated in Table 1? How were groups matched (e.g., 1:1)?

**Response:** p-values were calculated based on t-tests for continuous variables and chi square tests for categorical variables. The groups were matched 1:1.

Page 9, Paragraph 2: More detail is needed about the Deyo Charlson Comorbidity.

**Response:** The Deyo Charlson comorbidity index was calculated by assigning a weight of 1, 2, 3 or 6 to comorbid conditions, depending on their associated risk of mortality, and totalling the score to give an overall predicted mortality. The table below shows the ratings used in our study. We have not included this in the manuscript. We could include it as appendix if the journal prefers it.

<table>
<thead>
<tr>
<th>Charlson Comorbidity Index Calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Disease</strong></td>
</tr>
<tr>
<td>Myocardial infarct</td>
</tr>
<tr>
<td>Congestive heart failure</td>
</tr>
<tr>
<td>Peripheral vascular disease</td>
</tr>
<tr>
<td>Cerebrovascular disease</td>
</tr>
<tr>
<td>Dementia</td>
</tr>
<tr>
<td>Chronic pulmonary disease</td>
</tr>
<tr>
<td>Connective tissue disease</td>
</tr>
<tr>
<td>Ulcer disease</td>
</tr>
<tr>
<td>Mild liver disease</td>
</tr>
<tr>
<td>Diabetes</td>
</tr>
<tr>
<td>Hemiplegia</td>
</tr>
</tbody>
</table>
The Charlson comorbidity index has been applied in claims database analyses in the past. Some example citations:

Page 9, Paragraph 2: Why was a gamma regression chosen? How was it estimated? Did the model fit the data?

**Response:** We chose gamma regression as a widely accepted approach for assessing health care cost data.

Generalized gamma regression – a common method to address the skewed distribution of cost data – was used to compare the total healthcare costs of patients treated with adjunctive aripiprazole and patients treated with adjunctive ziprasidone, olanzapine, quetiapine or risperidone.

Gamma regressions were used to compare outcomes because gamma distribution is suggested by many as a close approximation of cost data. For example, Diehr and colleagues compared different methods to model healthcare cost data and concluded that for understanding the effect of individual covariates on total costs, the gamma distribution might be preferred because it is a multiplicative model (Diehr 1999). This is also supported by other research work (Blough 1999;Bouckaert 1997;Gilmer 2004).


Page 9, Paragraph 2: Why were only those with healthcare costs in the follow-up period chosen? How could this have biased the results? Would a two-part model help address these concerns?

**Response:** We were interested in the effects of antipsychotic choice on post-treatment healthcare costs, adjusting for pre-index costs. This adjustment in addition to propensity score matching, which included pre-index psychiatric hospitalization, should adjust for bias reasonably well.

Page 13, Paragraph 2: The authors indicate, “These results suggest an economic advantage for aripiprazole based on the presence of incremental clinical benefit without higher total expenditures.” The cost analysis was not of average expenditures but of average expenditures given there was some spending. It is possible for the average to be higher even if the conditional average is lower.

**Response:** We made the a priori assumption that if a patient received an antipsychotic they must logically have positive costs (physician visits), and that patients without positive costs had questionable data. Therefore we only included patients that we believed had analyzable data.

Page 14, Paragraph 2: The authors indicate they did not address adding or switching antipsychotics. Yet, these factors could have been important to examine. Perhaps, they could compare the results of the two analyses – one analysis with only those who used the antipsychotic all the way through versus an analysis that includes those who switched and discontinued use. The authors assert, “Propensity score matching, however is a widely accepted method for minimizing the effects of treatment selection bias in observational data.” Instrumental variables (IV) is a widely accepted method for controlling unobserved selection bias. Future analyses using IV might seek to reduce the potential bias that remains in this analysis.

**Response:** Our study was limited to the initial choice of antipsychotic and censored patients who switched or added. This is a limitation and a follow-up study analyzing post-switching consequences would be informative. The design, however, was an inception rather than a simple cohort in order to better characterize initial treatment effects.

Instrumental variables are a newer approach to control for selection bias that are less widely accepted than propensity score matching primarily because
the selection of the IV can be controversial and often not available from available data.

Page 15, Paragraph 1: The authors conclude, “Total healthcare costs of patients treated with aripiprazole were lower than those with quetiapine.” This is not exactly accurate; overall healthcare costs were not studied.

**Response:** We studied all direct healthcare costs that were available in the claims dataset. Therefore we assert that total direct healthcare costs were studied.
Response to Reviewer 2: Angus Thompson

Major Compulsory Issues to be Addressed:

1. There are two concerns about the key statistical significance values reported in Table 3 (i.e. for “Comparator vs. Aripiprazole”).
   
   i. There are four paired-comparisons reported – each involving aripiprazole. Ordinarily, the p-level is adjusted to compensate for this form of multiple comparison. Thus, some of the marginal findings (i.e. for Olanzapine and Quetiapine) may not prove to be statistically significant.

   **Response:** Multivariate statistical modeling is not subject to the same requirements for adjustment in that “multiple comparisons” are not actually being made. This is more pertinent to an experimental study design that is subject to conventional power calculations. These four paired-comparisons in this study are conducted independently from each other. In statistics, the problems of the multiple comparisons occurs when one considers a set, or family, of statistical inferences simultaneously.


   ii. Most journals prefer significance levels to be set at p<0.01. Table 3 levels are set for 95% confidence intervals. Again, it is likely that some of the comparisons judged to be statistically significant will not prove to be so. In sum, considering these two points, it may be that none of the comparisons will prove to be statistically significant. At best, one might be marginally so.

   **Response:** Our pre-specified alpha was set at 0.05 (we stated that in the methods). This is the threshold of evidence that is typically accepted to avoid a type 2 error. While there is nothing “magical” about this threshold, it should also be understood that retrospective observational studies are inherently biased to the null hypothesis due to confounding by indication and the crudeness of the measurements. The probability that our results were due to chance may therefore be even lower than what is reflected by the confidence interval.

2. The propensity score-matching on mood stabilizers for Aripiprazole vs. Ziprasidone produced a higher percentage on all six of these mood stabilizers for Ziprasidone (Table 1). This may not look like much, but a simple binary (higher vs. lower) analysis results in a p level of about 0.03. Suggest either reselecting the sample, or explaining why my concern will not affect the results (note from the point just above that Ziprasidone showed the highest significance level).

   **Response:** Higher use of mood stabilizers in the ziprasidone sample would likely bias the results to the null hypothesis in that patients were more likely to
be on concomitant mood stabilizers and therefore “doubly” protected against hospitalization.

Furthermore, none of the p-values of six of these mood stabilizers for Ziprasidone versus Aripiprazole are significantly different (P values: 0.701, 0.787, 0.781, 0.623, 0.738, 0.932).

We calculated the average number of mood stabilizers in baseline for the comparison groups. The number of mood stabilizers for aripiprazole is close to the number of mood stabilizers for ziprasidone (0.68 vs. 0.71).

<table>
<thead>
<tr>
<th>average number of mood stabilizers in baseline</th>
<th>aripiprazole</th>
<th>ziprasidone</th>
</tr>
</thead>
<tbody>
<tr>
<td>matching aripiprazole to ziprasidone</td>
<td>pre Propensity Score matching</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>post Propensity Score matching</td>
<td>0.68</td>
</tr>
</tbody>
</table>

We did sensitivity analysis for the unmatched study sample, in which aripiprazole has higher number use of mood stabilizer than ziprasidone (0.78 versus 0.70). The adjusted hazard ratio of ziprasidone vs. aripiprazole is 1.571 (95% CI 1.195-2.066), p=0.0012.

3. The propensity-score matching exercise involved 59% of the 776 aripiprazole patient pool in the comparison with ziprasidone. The figures for the other three antipsychotics, however, were 95% or better. This means that the comparisons of aripiprazole with the other medications are based on two unequal samples (one for ziprasidone and one for the other three). Does this matter? It would be useful to allay the readers’ possible concerns about what this might mean.

Response: Propensity score matching is limited by the smallest sample size. With all comparators except ziprasidone, aripiprazole was the smaller sample. It should be noted that propensity score matching is designed to minimize differences between samples – in this case, on a 1:1 basis. It would be counterintuitive to expect that all four comparisons would use the exact same patients.

Discretionary Issues:
1. In spite of the attempts to control selection bias, studies on treatment seeking have shown us that adjustment for supposedly important covariates does not guarantee its (bias) removal. The authors should be commended for their approach to this issue and for mentioning the limitation. However, given the attention they have paid to the details it is surprising that they have not provided the readers with some conjectures about means for further improvement.
Response: We were hesitant to digress into methodological suggestions, although the nature of the analysis and incidental findings was very invigorating.

2. Why were only “positive healthcare costs” included in the economic analysis? Other forms of indirect costs are often illuminating. Please clarify this omission. If other costs could not be estimated, please say so and acknowledge what might be lost.

Response: “positive healthcare costs” are the only type that can be obtained from the dataset used, i.e. insurance claims. We excluded patients with non-positive costs (“negative” costs or zero costs) because those were likely unreliable data, i.e. if patients had prescriptions, they had to receive services that were billed. If such data were missing, then the entire patient file was deemed unreliable.

3. Would it have been useful/possible to assign medication costs to either psychiatric and medical treatment as was done for inpatient and outpatient Rx?

Response: We were hesitant to increase the granularity of the analysis, which was already rather extensive. This is one limitation, although the relative pharmacy costs seem to be reflective of the relative costs of the antipsychotics studied. This seems to indicate that the main pharmacy cost driver was the antipsychotic medication.

4. Why was aripiprazole selected as the reference antipsychotic? Was it selected on the same basis in the previous study on the same antipsychotics?

Response: The unadjusted psychiatric hospitalization rates were lowest for aripiprazole, as reflected in Table 2. It was therefore logical to use aripiprazole as the reference. Alternatively, we could have chosen ziprasidone, which had the highest rates, but that would have required propensity score matching with the smallest treatment group and unnecessarily sacrificing sample size.

5. A strength of the Cox proportional hazards analysis is that it makes allowance for drop-outs. But the ability to produce a number is one thing, the interpretation of what it means in light of the huge number censured here, is another. What does it mean to attribute an outcome to, for example, aripiprazole, when 88.3% of the sample discontinued its use (including those that switched to another antipsychotic) before the end of follow-up? At best, the comparisons pertain to the drug prescribed at time of the beginning of the study period – but many things occurred after that. Aripiprazole showed the highest rate of discontinuance – perhaps that accounts for the apparent benefit.

Response: This is an important point. Our interest in the results is that the choice of initial antipsychotic appears to have long-term consequences. Thus, the sequence of treatments is important since early decisions may
affect costs for a protracted period of time. The question of what, if any, other antipsychotics were received subsequent to the initial treatment is of interest, but would require a different study design.

6. It appears that psychiatric, medical, and pharmacy costs do not follow the same pattern across the atypical antipsychotic medications. If this is systematic, then there is something to be learned by examining the relationships further. The authors may wish to comment on this.

**Response:** We have described the patterns of psychiatric, medical and pharmacy costs across the atypical antipsychotic medications in the results. “Total mental health costs were lower for aripiprazole compared to ziprasidone and quetiapine but not different compared to olanzapine and risperidone. Compared to aripiprazole, total medical costs were higher for quetiapine but not different for all other comparators. Pharmacy costs were lower for olanzapine, risperidone and quetiapine and not different for ziprasidone. Total healthcare costs in the follow-up period were significantly lower for aripiprazole than quetiapine and not significantly different for the other comparators.” Further exploration of the relationship is beyond the study scope.