To the Editors,

We thank the reviewers for their thoughtful review of our science. We have addressed concerns comprehensively in a point-by-point fashion, making changes and edits where appropriate. We have included both a highlighted version with changes annotated and a clean copy of the text for your review.

Editor’s Comments

1. “Is there any overlap in included patients between the database used in this study and other available databases? Are the databases exclusive of one another?”

There are a number of administrative claims based datasets that can be potentially used for a similar analysis depending on adequate sample size. These include Medicare claims, Marketscan, and IMS Lifelink among others. There have already been published studies looking at the safety of novel anticoagulants using Medicare claims. While there have been no published studies using the other two datasets, there is NO overlap in data between Optum Labs data (used in this study) and Medicare, Marketscan or IMS Lifelink. The only available dataset where there would be some overlap with the Optum Labs data is the Health Care Cost Institute (HCCI) database. However, to our knowledge, there are no studies assessing outcomes/safety of any kind for any treatment that have been published using the HCCI database yet.

2. Request for provision of comparative stroke risk (i.e., safety). Per Dr. Burch: “Perform the analyses and place them in an appendix. A mention of the appendix could be put in the paper in the appropriate places, and the paper would not have to be revised much further than that with regard to this question.”

We have conducted an analysis comparing the stroke rates across treatments. The methods and results of this analysis have been added to the appendix (Table A4) per our discussion with the editor. The results of this analysis are presented in the table below.

Table. Events and Adjusted Hazards of Stroke Events

<table>
<thead>
<tr>
<th>Atrial Fibrillation</th>
<th>Dabigatran vs Warfarin Stroke</th>
<th>Rivaroxaban vs Warfarin Stroke</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events per 100 Patient-Yrs. (95% CI)</td>
<td>Hazard Ratio* (95% CI)</td>
<td>Hazard Ratio* (95% CI)</td>
</tr>
<tr>
<td>Dabigatran</td>
<td>Warfarin</td>
<td></td>
</tr>
<tr>
<td>Total Stroke Events</td>
<td>0.74 (0.53, 1.05)</td>
<td>0.72 (0.51, 1.01)</td>
</tr>
</tbody>
</table>

*Hazard Ratio Adjusted for those covariates significant at the <0.05 level.
Statistical Reviewer’s comments

3. “Abstract - The data source quoted here, although not incorrect, is misleading. The total population on which this analysis was based on is less than 20k individuals for each comparison. Please correct/clarify.”

Abstract has been revised, as requested.

4. “Abstract - Results only provide crude incidence rates of two of the groups (not for Warfarin) and no HRs. Please give incidence rates for all groups with confidence intervals (not currently provided) as well as the HRs (with CIs) for the main comparisons – Total bleeding by NOAC by AF/Non_AF. I assume that the age interaction was part of the primary objective and hence at least the most relevant HRs (e.g. for age 75+) should also appear in the Results so that the current “Conclusion” is adequately supported.

Abstract has been revised, as requested. Provision of this data does put the word count beyond that which is requested by the journal. If editor believes this data can be removed to accommodate the word count, we would be happy to comply.

5. “Related to the issue of total sample size, the 1:1 matching seems excessive as many of the warfarin cases are not used (more than 50% lost). Given that in some of these comparisons there are borderline statistically significant findings (“less GIB found in the dabigatran group”) it would have made sense have tried for as many matches as possible. This needs to be mentioned in the discussion.”

The reviewer identifies an important issue to consider in the context of propensity score matching, especially given the differences in sample sizes for our treatments. We have included the total sample available for warfarin patients in Figure 1 and in the results section- including the proportion that were matched. Prior research suggests that 1:many matching for propensity scores has the potential to increase precision with an increase in bias. (1) In other research looking at optimal matches (for untreated subjects) where 1:many matching is possible, the researchers suggest that the optimal matching is either 1:1 or 1:2. (2) Based on these recommendations, we decided on 1:1 matching approach. While it is possible, that a 1:2 approach might increase precision, it may also decrease the quality of the matches. We would be happy to revisit this if the editors and reviewers felt strongly about the approach. We have amended the discussion (pg.15) to reflect this point.
6. “In terms of reporting, the phrase “dabigatran had a slightly lower risk of GIB” (Page 10, line 5) needs changing. Based on the 95%CI there is no evidence of difference between the risks in the two groups.”
   This has been done, as requested.

7. “Figures 2 and 3: Please delete the multiple comparisons with the p-values associated (presented as text in the graphs). These would need to be adjusted (bonferroni?) to deal with multiple comparisons and in any case all the information is provided by the graph with the point estimate and confidence intervals. It would be more useful to instead provide the total sample for each of the Age groups by intervention.”
   We have updated figures 2, 3, A4 and A5, removing the multiple comparisons with the associated p-values (we had used Bonferroni for comparisons), and further we have updated the figures to include total sample for each age group.

8. “New data: The data for Figures A7 to A10 in the Appendix are not used/discussed in the manuscript. There seems to be a considerable difference between these graphs (A7 to A10) and what appears in Table 3 and Figures 1 and 2. Going through the “response to reviewers” it is apparent that these are the crude event rates by age group used to create Figures 2 and 3. Please add explanation to these in the Appendix.”
   Figures 2, 3, A4 and A5 demonstrate the interaction effect as a function of age and oral anticoagulant treatment. Figures A7 through A10, are unadjusted event rates by age group, requested by the prior reviewers. Clarifying text added to figures A7 through A10 in the appendix.

We hope that these changes and comments satisfy the residual concerns of the Editor and the statistical reviewer. On behalf of the authors, please allow me to thank you for your consideration of our work.


2. Austin PC. Statistical criteria for selecting the optimal number of untreated subjects matched to each treated subject when using many-to-one matching on the propensity score. Am J Epidemiol. [Research Support, Non-U.S. Gov't]. 2010 Nov 1;172(9):1092-7.