Dear Dr. Villanueva and colleagues,

Thank you kindly for your careful review of our manuscript and for the opportunity to submit a revised version for your consideration.

We provide below a point-by-point response to comments by reviewer 1 and the statistical editor. Two copies of the manuscript are provided, one with changes highlighted and the other a clean copy. All mention of line numbers in the point-by-point response refer to the clean version of the manuscript.

Sincerely,

Björn Pasternak

Reviewer: 1

The revision is much improved, and I thank the authors for responding to my comments and changing the paper. Some final recommendations.

Abstract: "Fluoroquinolone use was associated …" change to "After propensity score matching, Fluoroquinolone use was associated …"

Changed to "In a propensity score-matched cohort, fluoroquinolone use was associated…" (page 3, line 14)

Abstract: The 60-day window is added well to the paper, thank you. Also change here: "with an estimated absolute difference of 82 (95% CI, 15 to 181) cases of aortic aneurysm or dissection BY 60 DAYS per 1 million treatment episodes".

Changed as suggested

Methods: "The absolute risk difference for the 60-day risk period was estimated as …", please change to "The absolute RATE difference for the 60-day risk period was estimated as …" – as the equation provided gives a difference in rates, not risks, and hence the confusion in my earlier comment about clarifying this equation.

Changed as suggested

Results: Mention in the results section about whether there was evidence of non-proportional hazards in the 60 day period. This is currently given in the methods section.

Now moved from methods to results (page 11, line 30).

Results: "There was no increased risk of aortic aneurysm or dissection associated with fluoroquinolone exposure in the period of 61-120 days from start of treatment (hazard ratio 0.67; 95% CI, 0.40 to 1.11)." Here, the authors must also mention whether there was any strong evidence of non-proportional hazards between the earlier period upto 60 days and this later period. You cannot give this strong statement without comparison to the earlier period of 0 to 60 days. Indeed, it seems interesting that the CI for the later period has an upper bound of 1.11, and the lower bound of the CI for the earlier period is 1.12. Also, this is a secondary analysis according to the methods section, and I think should be labelled as such in the results (other secondary analyses are).

By stating that the cited sentence is a strong statement and by discussing a non-overlap between CIs of the primary 60-day period and the secondary 61-120-day period, the statistical editor appears to imply that we have declared that there was a significant difference between the primary 60-day period and the secondary 61-120-day period.

No such declaration is being made, and there was neither intention of making such a comparison statistically nor reason for doing so.

As suggested, the 61-120-day period is now explicitly labeled as a secondary analysis in the results section (page 12, line 3).

Results: I still do not understand the rationale for the dichotomisation of age at 64 years. The authors suggest this is common practice and that the value is near the mean, and was pre-specified. None of this is justification for a wrong analysis. If age is truly of interest, it is sub-standard to analyse it as a binary variable. There are many references for this. E.g. see [1-3] and the references therein. However, as this is a secondary analysis, I simply want to state that I disagree with the response, but that it is not a major stumbling block before publication. Perhaps a comment could be added to the Discussion.

Thank you.

Figure 4: Personally, I prefer to see the probability on the y axis (going from 0 to 1), and not the % (which goes from 0 to 100). With such small numbers, I think the reader may wrongly interpret the provided %s as probabilities.

Changed to 0-1 scale.

Figure 4 nicely shows that the risk of aortic aneurysm or dissection was extremely low in both groups. A very important addition. I think the authors should also give the cumulative incidence at 60 days in the text itself, for both groups, to emphasise this further. Perhaps after the provided number of cases per 1 million treatment episodes, as it gives (as the authors note to me) the same result of an additional 82 cases by 60 days per 1 million episodes. That is, say that using Figure 4 the cumulative incidence for the Fluoroquinolone and amoxicillin groups are (approximately for me) 0.020/100 = 0.0002 and 0.012/100 = 0.00012 , respectively. Therefore, the difference in risk of aortic aneurysm or dissection by 60 days is 0.00008. This corresponds to an extra 80 cases per 1 million episodes. (NB obviously use exact values unavailable to me, and give CI).

The cumulative incidence has now been introduced in the text, as below (page 11, line 43). We have chosen not to report the cumulative incidence difference between groups in order to keep the results to one single measure of the absolute difference (i.e. number of cases per 1 million treatment episodes):

"In the 60-day risk period, there were 64 cases of aortic aneurysm or dissection among 360,088 treatment episodes of fluoroquinolone use (incidence rate 1.2 per 1000 person-years), as compared with 40 cases among 360,088 treatment episodes of amoxicillin use (incidence rate 0.7 per 1000 person-years). The cumulative incidence of aortic aneurysm or dissection at 60 days was 0.00020 for episodes of fluoroquinolone use and 0.00012 for episodes of amoxicillin use (Figure 2). There was an increased risk of aortic aneurysm or dissection associated with fluoroquinolone use (hazard ratio 1.66; 95% CI,

1.12 to 2.46). This corresponded to an absolute difference of 82 (95% CI, 15 to 181) cases of aortic aneurysm or dissection per 1 million treatment episodes in the 60-day risk period ~~(Figure 2)~~."

Note: The old Figure 4 has been moved to earlier in the text and therefore renumbered; that figure is the new Figure 2.

Figure 2 should be removed. The bars adds nothing to the result at the top of the figure, which should be in the text.

(The old) Figure 2 has now been removed.

Figure 3: again what do the bars tell us? The key results are the number of cases, not the size of the bars. And the axis scale dictates whether one bar looks big relative to another bar. I think the results are better given in a table. Also, no of cases is the y axis. But out of how many people in each group? One group has 5 times more participants.

We are not completely certain with respect to the entire meaning of this comment.

a) As is clear from methods, results text, Figure 1, and Table 1, this is a 1:1-matched cohort with 360,088 episodes included each exposure group. The statement that one group has 5 times more participants is incorrect.
b) The bars do show the number of cases.

Figure 3 has now been converted to a new Table 3.

Discussion: "methods to minimize the possibility of confounding" – I think this is too strong. I would suggest 'methods to adjust for 47 potential confounding factors". Similarly, end of page 43.

This summary statement not only refers to propensity score-matching but also to the active comparator design, which is key to control for confounding. Hence, limiting this statement to "adjust for 47 potential confounding factors" would not give the whole picture.
Wording has been softened up as follows:

Page 14, line 24:
"…methods to attempt to limit ~~minimize~~ the possibility of confounding,.."

Page 16, line 9:
"An important concern in any observational study is the possibility of confounding. To attempt to limit ~~minimize~~ this possibility, we used an active comparator to limit confounding by factors associated with filling an antibiotic prescription, including confounding by indication, and propensity score-matching derived from a range of covariates.".