We thank again the reviewer for these constructive comments, which helped us to further improve the paper. Our responses to all the comments are listed below, written in *red italic text*. The adapted text in the paper is indicated by green background colour. (However, we have not indicated minor changes related to grammar or spelling.)

Reviewer 1

The authors have improved the manuscript but there are still several typos and errors that need to be cleaned up.

Thank you, we have carefully reviewed the paper and removed all typos found.

Also, there are still several sweeping statements and loose logic throughout the paper that needs it be tighten or qualified. Please check through the paper carefully, especially the introduction, discussion and conclusion.

We apologize for these oversights; we tried to qualify a few in the mentioned places.

For example, “The reliability of these ratios was confirmed by comparing them” – confirmed is too strong. First, your estimates are only consistent with theirs and relative comparison cannot confirm that you are right. Second, you are also assuming that their estimates are correct. I suggest that you tone down such claims (throughout the paper). I suggest something like “Our estimates are consistent with …., which provides (some) support that they are (maybe) reliable.

Thank you for pointing this out, we have removed the indicated sentence and inserted a subsequent sentence (lines 374) with what was suggested: "The ratios in Table 4 (number of casualties to the number of crashes) are in line with the ratios used in road crash cost studies in other African countries, as shown in Table 5, which provides some support for the reliability of our estimates"

Another example is “The study demonstrates that under-reporting” – again, demonstrates is too strong. It implies that you have direct evidence from your data and analysis or concrete evidence but the evidence provided are indirect and circumstantial although very indicative.

Thank you, we agree the word demonstrate is overconfident in this context so we replaced the verb "demonstrate" with "indicate".

There are still several key assumptions besides the using the average that need to be clearly stated and discussed. I suggest checking each step of the calculation. Discussion should include why the assumption is made or why this the method chosen, as well as the possible sources of bias and how it would affect the estimates (under-estimate/over-estimate).

For example, the main source of data on fatality under-reporting for the study is the WHO. You are assuming that WHO estimates are correct. If I remember correctly, the estimates are obtained from a Poison regression using a cross sectional data from different countries and the accuracy is not that high, which is expected as it does not include country-specific data but uses only very generic across the board information. The confidence intervals are therefore quite large. The actual estimated value thus may be much higher or much lower.

Thanks for this comment. WHO uses a negative binomial regression model to estimate road fatalities in countries without eligible death registration data. The model is based on variables theoretically relevant as predictors of the number of road fatalities like GDP, population, working population, health system access, vehicles per capita, etc.

In this study, in the absence of other information or other useful data sources, the WHO estimated value is taken as a reference. In the text we added evidence supporting this choice.

A similar example is the assumption that the hospital survey in Guinea is a good estimate for the under-reporting rate in Cameroon. Please justify and discusses its implications on your estimates.

Thanks, we provided justification and added some evidence supporting our choice. We have also indicated in the revised version that we are making reference to the pattern observed (ratios) in the hospital data in Guinea and not the level of under-reporting.

Is pain, grieve and suffering limited to production loss? We all suffer from these in any injuries, regardless of whether we are contributing to production or not. Are you assuming that children and the elderly do not suffer from pain and grieve? Do we not grieve for our loved ones who are injured or killed? Do we not take leave from work because our loved ones are injured or killed? How about inconvenience, time cost and other intangible costs? If my car is damage and under repair, don't I have to suffer from these non-out-of-pocket costs? This study deals mainly with explicit costs and ignores implicit and intangible costs. This assumption should be clearly stated at the beginning of the methods sections. The implications on the estimates should also be discussed in the results or discussion section.

Thank you for making this comment, however we would like to point out that our approach also considers human costs, which include the intangible costs of quality of life. Ideally, the WTP approach is aimed at calculating human costs, since there is no market price for such intangible losses. However, since a country-specific WTP-study has not been undertaken being complex and resource demanding, human costs have been addressed considering evidence from other countries. In our study, a value transfer function developed by a World Bank study that relate the Value of a Statistical Life (VSL) to GDP per capita was adopted.

Assumptions about data also need to be stated and discussed. Besides police and hospital data, there are other sources such as insurance data. What are the relative strengths and weaknesses of the different sources and what are the implications on the estimates (over or under) for the data source you chosen?

Thanks for this comment. We added a paragraph in the discussion part referring to the main data sources used in the paper (Table A13 of the annex)