PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Rates of adverse events of antiretroviral therapy in women living with HIV/AIDS: a systematic review and meta-analysis
AUTHORS	de Oliveira, Jardel; Alves, Maíra; Lopes, Luis; Motter, Fabiane; Iwami, Rodrigo; Bergamaschi, Cristiane; Silva, Marcus; Scalco, Diogo; Lucio, Donavan; Mazzei, Lauren; Derech, Rodrigo; Itria, Alexander; Barreto, Jorge; Lopes, Luciane

VERSION 1 – REVIEW

REVIEWER	Ritah Mutagonda Muhimbili University of Health and Allied Sciences, Clinical Pharmacy and Pharmacology
REVIEW RETURNED	24-Sep-2023

GENERAL COMMENTS	This is a very important manuscript that will add to the body of knowledge on the incidence of ADEs of ART among women living with HIV/AIDS. There are a few observations for consideration: Abstract: 1. In the objective, the sentence can be rephrased. Example: Objective: This study/review aimed to describe the incidence ..... 2. In the methodology, it would be very informative if the inclusion years in which the randomized controlled trials were reviewed were described. The review included publications from which year? 3. In the results, 10 studies met the eligibility, and nine studies were assessed as being at high bias, so is this the review of 1 study? 4. In the results, the first information I would expect to be presented is the overall mean incidence rate of adverse events, followed by other information such as treatment discontinuation, etc. 5. The conclusion should be based on the main objective of the review, which was to describe the incidence of AE of ART in women living with HIV/AIDs. The scarcity of information can be described as the limitation of this study. Introduction: 1. The data provided in the first paragraph does not have references. 2. I suggest you use 'HIV/AIDS standard treatment guidelines' rather than protocols in the second paragraph. 3. The categorization of serious events described under the outcomes is not clear. How did the author distinguish grade 4 events from serious events? I think the worldwide standard categorization used by FDAs should be used to categorize ADEs to make sense of the results of this work. 4. How death was treated in this study is not clearly explained.

	Methodology: 1. In the first paragraph, it is better to describe that objective rather than stating objective 2 so that even the person who has not read the previous publication will understand what objective 2 was. Results 1. See the previous comment related to the categorization of ADEs. Discussion 1. The summary of the main results should be in line with your objective. What was the incidence rate of ADEs in women living with HIV/AIDS on ART regimens? On reproductive system (none reported) on bones (6.15 events per 1000 person years), SAE? Regimen with most events? Death? 2. The entire discussion should be revised. Implications for research The authors should start by stating briefly the significant findings of this work and then their implications and recommendations. Conclusion This part should show your major findings before stating limitations and recommendations.

REVIEWER	Bishara J Freij Beaumont Children's Hospital, Pediatrics
REVIEW RETURNED	27-Nov-2023

GENERAL COMMENTS	Drs. de Oliveira and colleagues have done a formidable job in putting this paper together, along with very thorough supplemental materials. The conclusions are appropriate in that there is a dearth of data focusing on women with HIV and their different biology as it relates to the infection and its treatments. The comments below are not meant to detract from this paper but to highlight the available data limitations even more: 1) While the study included all women who received at least one dose of a drug, it is not clear to me what the compliance rates were since adverse drug effects may be less prominent if you take the drug less often than recommended. 2) The health status of women (not just their CD4+ count/percentage) with adverse effects is not defined in terms of whether or not they had co-morbidities. For example, in the study of Firnhaber, many women had tuberculosis (pulmonary and extra-pulmonary). How many were on anti-tuberculous multi-drug regimens which can impact observed clinical and laboratory adverse effects? 3) The longest duration of treatment in the analyzed studies was 184 weeks (Firnhaber, 2015). Since HIV has been transformed into a chronic disease with life-long treatment, this is hardly enough time to recognize all long-term adverse effects, including drug-related cognitive decline or progressive cardiac ill effects, bone density drops, etc. 4) Pregnant women on HV medications will generally continue the treatment during pregnancy. Given the differences in drug pharmacokinetics and disposition in this population, compliance problems with morning sickness, the presence of a fetus to consider, etc., their adverse effect profiles should be evaluated as well (but not included with non-pregnant women). Similarly for women choosing to breastfeed.

	5) Another unaddressed issue is the ageing of women with HIV and how their ART-related adverse effects might be different than those among younger HIV-infected women. In the included studies, the mean or median ages are from the upper 20s to the mid-30s. 6) The authors state that they excluded 110 studies because they did not "describe adverse effects of interest for this review". I am not sure what that means and would appreciate a clearer statement.

REVIEWER	John Bosco Matovu ICAP at Columbia University, CQUIN
REVIEW RETURNED	03-Dec-2023

GENERAL COMMENTS	Through a meta-analysis, authors have explored an interesting topic of determining the incidence of adverse events among women who take ART, and findings have revealed that this information is scarce. The statistical methods are comprehensive and sound! A few comments have been included in the PDF copy, some are semantics, and some sentences may need rephrasing and these constitute minor changes The reviewer provided a marked copy with additional comments. Please contact the publisher for full details.

REVIEWER	Luojia Deng Shanghai Jiao Tong University, Department of Bioinformatics and Biostatistics
REVIEW RETURNED	14-Dec-2023

GENERAL COMMENTS	The manuscript systematically reviewed and analyzed the incidence rate of adverse events (AE) and antiretroviral therapy (ART) in women living with HIV/AIDS. The authors collected a wide array of ART regimens and adverse events, which is a heavy workload. And since the included articles are all RCTs, the quality of the evidence is also high. Below are some of my comments. 1. Page 7, Lines 7-10: “We excluded trials involving pregnant or breastfeeding women or with a focus on co-infection with tuberculosis, hepatitis B or C”. I appreciate this approach, as indeed, co-infections and other comorbidities can affect AE with ART. By excluding such studies, it helps to focus more on the impact of ART itself on AEs. 2. Page 8, Lines 10-19. Two Excel spreadsheets were provided here, and upon inspection, I noticed that may entries are Portuguese. It would be advisable to standardized these to English. The same issue is also present in the supplementary file. The authors should carefully review and modify these documents to ensure uniformity in language. 3. Page 8, “Risk of Bias” section. Using the RoB2 tools to assess the risk of bias in RCT studies is appropriate. However, I disagree with the statement about publication bias. Despite the manuscript calculating the incidence rate of AE in a single group, rather than a ratio or difference between two groups, it is still possible to conduct tests for publication bias, such as funnel plots and Begg’s test. Since the authors have used funnel plots, this statement should be revised. 4. Page 8, “Statistical analysis” section. Unlike frequentist approach, the authors have employed a Bayesian random-effects model to pool the incidence rates and used 95% predictive intervals

	to show between-group heterogeneity. A more conventional metric like $I2$ might be more intuitive for readers to understand the degree of heterogeneity, as PI makes it difficult to directly discern whether heterogeneity is high or low. In addition, I suggest reorganizing and rewriting the 'statistical analysis' section. The current format, with one-sentence paragraphs, does not offer a good reading experience. 5. When abbreviation CrI (credibility interval) first appears in the text, it should be linked to the term "credibility interval" like 'credibility interval (CrI)'. I could not find a corresponding explanation in the text, which was confusing during my first reading. 6. In the result section, where incidence rates have been pooled according to various outcomes, it might be beneficial to conduct subgroup analyses based on population characteristics if possible. This approach could provide more nuanced insights. 7. Last but not least, is there a specific definition for the AEs included in the study? I might have missed it, but I couldn't find a detailed description of what constitutes an AE. A table or a section in the text listing these events would be helpful. While Table S3 mentions some, it is unclear if these are all the events considers. For example, for clinical/laboratory-defined AEs. What are the exact criteria? These details should be clearly specified.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Ritah Mutagonda, Muhimbili University of Health and Allied Sciences

Comments to the Author:
This is a very important manuscript that will add to the body of knowledge on the incidence of ADEs of ART among women living with HIV/AIDS. There are a few observations for consideration:

Abstract:
1. In the objective, the sentence can be rephrased. Example: Objective: This study/review aimed to describe the incidence .....
Thank you. We adjust as suggested.

2. In the methodology, it would be very informative if the inclusion years in which the randomized controlled trials were reviewed were described. The review included publications from which year?
Thank you. We restructured the abstract as suggested by the editor and included this information.

3. In the results, 10 studies met the eligibility, and nine studies were assessed as being at high bias, so is this the review of 1 study?
Thank you. We restructured the abstract as suggested by the editor and we think this has become clearer.

4. In the results, the first information I would expect to be presented is the overall mean incidence rate of adverse events, followed by other information such as treatment discontinuation, etc.
Thank you. We change this.

5. The conclusion should be based on the main objective of the review, which was to describe the incidence of AE of ART in women living with HIV/AIDs. The scarcity of information can be described as the limitation of this study.
Thank you. We change the conclusion.

Introduction:

1. The data provided in the first paragraph does not have references.

These data are in reference 1, placed at the end of the paragraph.

2. I suggest you use 'HIV/AIDS standard treatment guidelines' rather than protocols in the second paragraph.

Thank you. We change this.

3. The categorization of serious events described under the outcomes is not clear. How did the author distinguish grade 4 events from serious events? I think the worldwide standard categorization used by FDAs should be used to categorize ADEs to make sense of the results of this work.

Thank you. When we planned our review, we looked to the FDA's definition of serious adverse events and the classification of grade 3 and 4 adverse events from the Division of AIDS of the U.S. Department of Health and Human Services, National Institutes of Health, National Institute of Allergy and Infectious Diseases (DAIDS Table for Grading the Severity of Adult and Pediatric Adverse Events). We understand your point of view and would like to classify by these definitions. However, we separated adverse events into grades 3 or 4, grade 3, grade 4, and serious, according to the definition used by the authors of the studies, as described in the methods. We consider that it would be an error for us reviewers to classify adverse events differently than the authors, since we were not directly collecting this data from study participants and the details given in the articles would not allow us to do this reclassification. In any case, seven studies classified the gravity (grade) of adverse events using the DAIDS. The others did not describe which source they used for this classification. We include this in the description of study characteristics in the results.

4. How death was treated in this study is not clearly explained.

We appreciate the reviewer for raising this important aspect. Our estimates indicate that all-cause mortality was low across the included studies, with an average of 4.47 deaths per 1000 person- years. Therefore, we do not expect a major impact of death on the estimation of incidence rates from non-mortality outcomes. However, in our Poisson models, we explicitly accounted for participants lost to follow-up (including those who died) by adjusting the denominator, the total follow-up in person-years. In our analyses, all patients who died or dropped out for any reason contributed to half of the follow-up time. From a statistical point of view, this is a reasonable assumption, since we expected deaths/losses to follow up due to any cause to be scattered across the follow-up period.

To address the reviewer's remark, we amended the statistical methods section as follows:

"When the actual total time at risk was not reported, we approximated the total person-years at risk using the mean follow-up duration, accounting for the all-cause losses to follow-up or withdrawals. We multiplied the reported mean follow-up by the number of participants analyzed, subtracting from it the product of half the mean follow-up and the number of participants not included in the analyses (all-cause dropouts/withdrawals)."

Methodology:

1. In the first paragraph, it is better to describe that objective rather than stating objective 2 so that even the person who has not read the previous publication will understand what objective 2 was.

Thank you. We adjust as suggested.

Results

1. See the previous comment related to the categorization of ADEs.

We modified the sequence of presentation of results, starting with the primary outcomes (total adverse events and discontinuation of treatment) and then the secondary ones.

Discussion

1. The summary of the main results should be in line with your objective. What was the incidence rate of ADEs in women living with HIV/AIDS on ART regimens? On reproductive system (none reported) on bones (6.15 events per 1000 person years), SAE? Regimen with most events? Death?

Thank you. We modified the topic 'summary of the main results' and started with the main results of the primary outcomes. We believe it is not necessary to repeat the other results, given the limitation of the evidence found. We add that the certainty of evidence was very low for all outcomes evaluated. Outcomes on bone health, treatment discontinuation and death were cited below, in the topic 'comparison with other reviews', as there was data in the literature for discussion. We understand that we should not highlight results on the ARV regimen with more adverse events or deaths, given the limitations found. Therefore, we prefer the phrase we previously mentioned in the discussion: "We could not establish recommendations on a preferential ARV regimen for use in women with HIV/AIDS, aiming to minimize the risk of adverse events and improve treatment adherence."

2. The entire discussion should be revised.

Sorry, we don't understand your request. The discussion is organized into the following topics: summary of main results; comparison with other reviews; strengths and weaknesses; implications for research; clinical implications. Evidence is weak to highlight specific results.

Implications for research

The authors should start by stating briefly the significant findings of this work and then their implications and recommendations.

We made the changes mentioned above.

Conclusion

This part should show your major findings before stating limitations and recommendations.

Thank you. We have added the main results in the discussion and partially reformulated the conclusion.

Reviewer: 2

Dr. Bishara J Freij, Beaumont Children's Hospital

Comments to the Author:

Drs. de Oliveira and colleagues have done a formidable job in putting this paper together, along with very thorough supplemental materials. The conclusions are appropriate in that there is a dearth of data focusing on women with HIV and their different biology as it relates to the infection and its treatments.

The comments below are not meant to detract from this paper but to highlight the available data limitations even more:

1) While the study included all women who received at least one dose of a drug, it is not clear to me what the compliance rates were since adverse drug effects may be less prominent if you take the drug less often than recommended.

Thank you. We agree with that. Three of the studies did not describe adherence to treatment. The others considered adherence by counting the difference between the total number of pills supplied and the number of pills not taken (returned), divided by the total number of pills delivered. A high level of adherence was usually considered when at least 95% of expected doses were taken. One of the

studies reported adherence by region in which the study was conducted, ranging from 61% to 92%. In the others, adherence varied from 81% to 99.7%. None of the studies presented the results of adverse events according to percentage or level of adherence to treatment. We include this as a limitation in the discussion.

2) The health status of women (not just their CD4+ count/percentage) with adverse effects is not defined in terms of whether or not they had co-morbidities. For example, in the study of Firnhaber, many women had tuberculosis (pulmonary and extra-pulmonary). How many were on anti-tuberculous multi-drug regimens which can impact observed clinical and laboratory adverse effects?
We describe the comorbidities presented in the studies in table 2. Firnhaber's study contributed only one outcome of interest to our review (laboratory adverse events). In this study, there were less than 10% of participants with tuberculosis, who were possibly receiving concomitant treatment for HIV and tuberculosis. In any case, the ARV regimens used in this study do not tend to cause an increase in adverse events resulting from interactions with the standard tuberculosis treatment regimen. What may occur is a reduction in the effect of efavirenz and zidovudine.

3) The longest duration of treatment in the analyzed studies was 184 weeks (Firnhaber, 2015). Since HIV has been transformed into a chronic disease with life-long treatment, this is hardly enough time to recognize all long-term adverse effects, including drug-related cognitive decline or progressive cardiac ill effects, bone density drops, etc.
Thank you. We include these limitations in the discussion.

4) Pregnant women on HIV medications will generally continue the treatment during pregnancy. Given the differences in drug pharmacokinetics and disposition in this population, compliance problems with morning sickness, the presence of a fetus to

consider, etc., their adverse effect profiles should be evaluated as well (but not included with non-pregnant women). Similarly for women choosing to breastfeed.
We agree with that. Therefore, we considered studies with pregnant women and women during breastfeeding as an exclusion criterion. Also, because this population of women has already been covered more in studies and has specific clinical practice guidelines. In general, the studies included in our review excluded pregnant women during participant selection and performed pregnancy tests periodically during follow-up. One study did not describe whether pregnancies occurred during follow-up and another only mentions that there was no statistical difference between the groups in terms of women who became pregnant during the study and that safety outcomes were not affected by pregnancies. In six other studies, the percentage of women who became pregnant was generally low, ranging from 0.5% to 4.1%. In studies by Lockman et al. this percentage was 5.8% (2010) and 9.4% (2021)

5) Another unaddressed issue is the ageing of women with HIV and how their ART-related adverse effects might be different than those among younger HIV-infected women. In the included studies, the mean or median ages are from the upper 20s to the mid-30s.
We agree with you. The average age in studies was what you mentioned. Because of this, we had put this into the discussion: “We found few data on elderly women and none of the studies included children.”

6) The authors state that they excluded 110 studies because they did not "describe adverse effects of interest for this review". I am not sure what that means and would appreciate a clearer statement.
Sorry, we don’t follow it. These were studies that did not describe any of the primary and secondary outcomes of interest for our review.

Reviewer: 3
Dr. John Bosco Matovu, ICAP at Columbia University Comments to the Author:
Through a meta-analysis, authors have explored an interesting topic of determining the incidence of adverse events among women who take ART, and findings have revealed that this information is scarce.
The statistical methods are comprehensive and sound!
A few comments have been included in the PDF copy, some are semantics, and some sentences may need rephrasing and these constitute minor changes.
Thank you. We adjusted the objective and the acronym for credibility interval in the abstract. We corrected some writing suggestions, but we did not understand others that were highlighted but without comment. We changed the order of paragraphs when describing outcomes in methods as suggested. We brought the last paragraph to the beginning, as it refers to the primary and secondary outcomes, and only then did we talk about the adverse events on women's bone and reproductive health (in the second paragraph).

Did you consider using Credible Intervals at 89%? 95% might not be the most appropriate for Bayesian posterior distributions, potentially lacking stability if not enough posterior samples are drawn-Kruschke, 2014 - JUST FOR CURIORITY

Thank you for your question. We used 95% credible intervals in all analyses. We employed three chains with 166,667 simulations each, totaling approximately 500,000 simulations. We carefully conducted model diagnostic checks. Our assessments revealed an excellent mix, close to zero autocorrelation, with no evidence for non-convergence for all models. All posterior distributions were reasonably symmetric, except for those from more uncommon outcomes, which were negatively skewed. However, upon careful assessment, we did not identify any cases in which high-density intervals would be needed and opted to use credible intervals based on percentiles, which is the traditional approach in Bayesian inference. Importantly, all analyses were rerun with a frequentist approach (multilevel Poisson random-effects model [see supplementary Figures S3 to S7]), and the results were compatible with the main analyses, indicating that the use of 95% credible intervals is appropriate in our analyses.

Reviewer: 4
Miss Luojia Deng, Shanghai Jiao Tong University Comments to the Author:
The manuscript systematically reviewed and analyzed the incidence rate of adverse events (AE) and antiretroviral therapy (ART) in women living with HIV/AIDS. The authors collected a wide array of ART regimens and adverse events, which is a heavy workload. And since the included articles are all RCTs, the quality of the evidence is also high. Below are some of my comments.
1. Page 7, Lines 7-10: “We excluded trials involving pregnant or breastfeeding women or with a focus on co-infection with tuberculosis, hepatitis B or C”. I appreciate this approach, as indeed, co-infections and other comorbidities can affect AE with ART. By excluding such studies, it helps to focus more on the impact of ART itself on AEs.
Thank you.

2. Page 8, Lines 10-19. Two Excel spreadsheets were provided here, and upon inspection, I noticed that may entries are Portuguese. It would be advisable to standardized these to English. The same issue is also present in the supplementary file. The authors should carefully review and modify these documents to ensure uniformity in language.
Thank you. We adjust the language.

3. Page 8, “Risk of Bias” section. Using the RoB2 tools to assess the risk of bias in RCT studies is appropriate. However, I disagree with the statement about publication bias. Despite the manuscript calculating the incidence rate of AE in a single group, rather than a ratio or difference between two groups, it is still possible to conduct tests for publication bias, such as funnel plots and Begg’s test. Since the authors have used funnel plots, this statement should be revised.

Thank you for your remark. We agree with the reviewer and provide readers with funnel plots for all outcomes with 10 or more estimates [1] (see supplementary figures S1 and S2). We did not create a funnel plot for adverse event-related mortality sparse data (too many studies with zero events). In addition, we did not use the Begg and Mazumdar rank correlation test because this method is no longer recommended [2]. We are not aware of any statistical test suitable for assessing funnel plots specifically for binomial and Poisson outcomes [3].

References

1. Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., ... & Higgins, J. P. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. Bmj, 343.

2. Sterne JAC, Egger M, Moher D, Boutron I (editors). Chapter 10: Addressing reporting biases. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors), Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017), Cochrane, 2017. Available from www.training.cochrane.org/handbook.

3. Hunter, J. P., Saratzis, A., Sutton, A. J., Boucher, R. H., Sayers, R. D., & Bown, M. J. (2014). In meta-analyses of proportion studies, funnel plots were found to be an inaccurate method of assessing publication bias. Journal of clinical epidemiology, 67(8), 897-903.

4. Page 8, “Statistical analysis” section. Unlike frequentist approach, the authors have employed a Bayesian random-effects model to pool the incidence rates and used 95% predictive intervals to show between-group heterogeneity. A more conventional metric like I2 might be more intuitive for readers to understand the degree of heterogeneity, as PI makes it difficult to directly discern whether heterogeneity is high or low. In addition, I suggest reorganizing and rewriting the ‘statistical analysis’ section. The current format, with one-sentence paragraphs, does not offer a good reading experience.

Thank you. We understand the reviewer's concern regarding the interpretability of the statistical heterogeneity. However, the I^2 is not an appropriate measure of heterogeneity when the within- study variances are tiny (e.g., as in large samples) [1-3]. The 95% prediction intervals are the most suitable approach to describe heterogeneity in a meta-analysis, regardless of the study sizes, because they describe the expected variation in true effects across 95% of the populations [4,5]. To convey a better reading experience and more information on how to interpret the 95% prediction intervals, we amended the statistical analysis section as follows:

"We quantified the between-study heterogeneity using 95% predictive intervals, which describe the expected variation in true incidence rates over different settings and populations."

We group the sentences in the 'statistical analysis' section.

References

1. Schwarzer, G., Schumacher, M., & Rücker, G. (2017). Sole reliance on I2 may mislead. Heart, 103(18), 1471-1472.

2. Rücker, G., Schwarzer, G., Carpenter, J. R., & Schumacher, M. (2008). Undue reliance on I 2 in assessing heterogeneity may mislead. BMC medical research methodology, 8, 1-9.

3. Borenstein, M. (2020). Research Note: In a meta-analysis, the I2 index does not tell us how much the effect size varies across studies. Journal of physiotherapy, 66(2), 135-139.

4. Borenstein, M. (2023). Avoiding common mistakes in meta-analysis: Understanding the distinct roles of Q, I-squared, tau-squared, and the prediction interval in reporting heterogeneity. Research Synthesis Methods.

5. IntHout, J., Ioannidis, J. P., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. BMJ open, 6(7), e010247.

5. When abbreviation CrI (credibility interval) first appears in the text, it should be linked to the term “credibility interval” like ‘credibility interval (CrI)’. I could not find a corresponding explanation in the text, which was confusing during my first reading.
Thank you. We adjust this.

6. In the result section, where incidence rates have been pooled according to various outcomes, it might be beneficial to conduct subgroup analyses based on population characteristics if possible. This approach could provide more nuanced insights.
The wide credibility intervals found in our review point to the great heterogeneity of the results presented and reduce the validity of possible subgroup analyses. That is why we chose to include these analyzes in the supplemental material.

7. Last but not least, is there a specific definition for the AEs included in the study? I might have missed it, but I couldn’t find a detailed description of what constitutes an AE. A table or a section in the text listing these events would be helpful. While Table S3 mentions some, it is unclear if these are all the events considers. For example, for clinical/laboratory-defined AEs. What are the exact criteria? These details should be clearly specified.
When we planned our review, we looked to the FDA's definition of serious adverse events and the classification of grade 3 and 4 adverse events from the Division of AIDS of the U.S. Department of Health and Human Services, National Institutes of Health, National Institute of Allergy and Infectious Diseases (DAIDS Table for Grading the Severity of Adult and Pediatric Adverse Events). We separated adverse events into grades 3 or 4, grade 3, grade 4, and serious, according to the definition used by the authors of the studies, as described in the methods. Seven studies classified the gravity (grade) of adverse events using the DAIDS. The others did not describe which source they used for this classification. We include this in the description of study characteristics in the results.

VERSION 2 – REVIEW

REVIEWER	Luojia Deng Shanghai Jiao Tong University, Department of Bioinformatics and Biostatistics
REVIEW RETURNED	15-May-2024

GENERAL COMMENTS	Most of my previous comments have been addressed in the response, but there are still some non-English phrases and sentences in the 'detailed statistical analysis' session of the supplementary file.