

Peer-Review comments and authors responses

“Effects of aerobic or resistance training on chronic inflammatory markers in postmenopausal women with overweight or obesity: A systematic review.”

Dear Reviewers,

Thank you for your thoughtful and constructive feedback on our manuscript. We appreciate the time and effort you dedicated to reviewing our work, and your suggestions have significantly improved the clarity and scientific value of our results and methods section.

We have addressed each one of your comments below:

Reviewer 1

1. **Comment:** "It would have been interesting if the authors had also included yoga and other forms of exercise. This may be discussed in the limitations."

Response: *We focused exclusively on aerobic and resistance training to maintain methodological homogeneity, since mind–body exercises (e.g., yoga, tai chi) act through distinct neuroendocrine and stress-related pathways rather than through direct metabolic or cardiovascular adaptations. Future reviews may integrate these modalities to provide a broader understanding of exercise-induced anti-inflammatory mechanisms. In response to your comment, we have provided this explanation in the “Implications for Practice and Future Research” section of the discussion. Excluding other exercises has also been added as a limitation (in the limitations section) that could affect the generalizability of the findings.*

2. **Comment:** “Sedentary Definition- It will be good to clearly define what the included studies considered to be a sedentary lifestyle and whether the definition was uniform across studies included.”

Response: *The included studies defined sedentary lifestyle differently. Variability in the operational definitions of “sedentary” across studies introduces heterogeneity that may compromise the comparability of findings in a systematic review. This has been added to the results section, under the sub-heading “Specific Differences Between Included Studies”, as follows: “Variations in selection criteria: The included studies defined sedentary lifestyle in different ways. Arsenault et al., 2009 defined women as sedentary if they did not engage in exercise lasting more than 20 minutes on three or more days per week and took fewer than 8,000 steps per day, as assessed over the course of one week. Brochu et al., 2009 classified women as sedentary if they performed less than two hours per week of structured exercise. Ryan et al., 2014 and Silverman et al., 2009 both defined sedentary women as those performing less than 20 minutes of aerobic exercise twice per week. Similarly, Stewart et al., 2010 defined sedentary women as exercising for less than 20 minutes on fewer than three days per week and taking fewer than 8,000 steps per day, assessed over one week. The variability in the operational definitions of “sedentary” across studies introduces heterogeneity that may affect the comparability of findings in this systematic review.”*

3. **Comment:** “Exercise Setting and adherence: Please mention the setting of the exercise (at home / supervised / unsupervised) in various studies and the adherence to the exercise regimen. This is crucial for interpreting results. Additionally, how was the adherence to dietary regimen? In case of poor adherence to both, it is likely to get non-significant results.”

Response: *We recognize the importance of this as exercise setting and adherence varied widely across studies. This has been added to the results section, under the sub-heading “Specific Differences Between*

Included Studies”, as follows: “ *Exercise adherence: Exercise setting and adherence varied widely across studies, influencing intervention quality and outcome validity. Studies with fully supervised or partially supervised programs—such as those by Arsenault, Stewart, and Imayama—reported higher adherence and lower dropout rates. In contrast, Brochu et al. reported low completion rates despite supervision, while studies incorporating home-based or self-directed components (e.g., Silverman, You) achieved moderate adherence through periodic supervision and monitoring tools. Poor adherence or irregular participation diminishes the training stimulus, increasing variability and reducing statistical power—ultimately increasing the likelihood that studies will report non-significant results even when a true effect exists. Dietary adherence was similar to exercise adherence in each study.*”

- 4. Comment:** Follow-up Period: Please mention the follow-up period for various studies [median or range] earlier on in the text for the readers to get a perspective of what is meant by “long term.”

Response: *This has been addressed in the first paragraph of the results section: “All the studies had a follow-up period of six months, except Imayama et al, 2012, which had a follow-up period of 12 months.”*

- 5. Comment:** Does the mini-review conclude that exercise has a short-term effect but not a long-term effect? The authors may hypothesize this is possibly due to issues related to adherence to exercise, given reported dropout rates (e.g., up to 25% dropouts in some studies).

Response: *Because our review included studies with longer exercise durations (up to six months), reduced compliance over time could have influenced the results. This has been explained earlier in the discussion: “Given our focus on trials lasting at least six months, lower adherence over time may have attenuated the observed effects. Nevertheless, this does not imply that exercise lacks benefit for postmenopausal women. Instead, the pattern across studies indicates that combining exercise with dietary modification may enhance anti-inflammatory outcomes through synergistic effects on metabolism and immune regulation.”*

- 6. Comment:** Tables 3-5: Please mention the actual p value if available for non-significant and significant columns. Please include CRP, IL-6 , TNF- α etc in the headings of the columns where their values are mentioned in the respective tables(intervention and control group). A common heading with subheadings as intervention and control group could work and be more explanatory.

Response: *Unfortunately, the actual p-values were not stated in the included studies; hence, we could not provide them. We have updated the tables with the proposed changes.*

- 7. Comment:** Baseline Differences: It seems there was a difference in baseline IL levels (Interleukin levels) in the two groups in the study by You et al. Please discuss the implications of this finding in the discussion.

Response: *We appreciate this important input. In You et al. (2004), CRP and IL-6 decreased more with diet + exercise than with diet alone, but baseline levels were already lower in the exercise group, suggesting regression to the mean rather than a true exercise effect. This has been explained in the first paragraph of the discussion, as follows: “However, CRP and IL-6 levels in that study were already lower at baseline in the diet-plus-exercise group, suggesting possible baseline imbalance and regression-to-the-mean effects that may overestimate the true exercise impact. The small sample size (n = 34) in You et al. further limits statistical power and generalizability.”*

- 8. Comment:** A small sample size ("You et al " as an example) can be a potential limitation. This was the only study which found a significant difference among groups. It will be good to discuss the possibility as to how the results could be a consequence of this small sample size. Additionally, were the other studies underpowered to detect a difference among groups.

Response: *We fully agree with your comment, You et. al. had a small sample size of 34 participants, and this could have affected the study's validity. This has been explained in the discussion: “Methodological Variations: Sample sizes varied substantially. Smaller sample sizes, such as those in You et al. (2004), can reduce statistical power, leading to underestimations of effects or an inability to detect actual changes.”*

Reviewer 2

- 1. Comment:** The authors chose the age of 50 as the minimum age criterion, although the age of 40 is considered the absolute threshold for the natural vs. primary ovarian insufficiency-associated menopauses. At the same time, authors mention that two studies had inclusion criteria of an age younger than 50. Therefore, revising the age threshold for defining menopause is recommended to avoid conflicting presentations. If the authors want to keep 50 as the threshold, it should be clinically justified; why not 49 or 51, but 50?

Response: *We fully agree with your suggestion to reconsider the age threshold. We decided to exclude the age range of 50, and keep only "post-menopausal women" as inclusion criteria. There were no other studies that fit the inclusion criteria with this adjustment, so we kept only the 7 studies already analyzed. Our population inclusion criteria now states: "(3) population: postmenopausal women with a body mass index (BMI) ≥ 25 kg/m², classified as overweight or obese".*

- 2. Comment:** In the results section, p-values from the previous study analyses are presented. Given that the statistical framework of the original article is not fully discussed here, the presentation of p-values alone doesn't add much scientific value beyond indicating that the findings were significant. Instead, it would be recommended to present the change in the effect sizes whenever applicable. For example, there was an increase in the X parameter by y-folds.

Response: *Ryan et al. (2014) and You et al. (2004) reported significant p-values, but no effect sizes were reported in the text or tables. We estimated the CRP effect size difference in Ryan et. al (2014) using Figure 1, and in You et al. (2004) by subtracting the before and after values for IL-6 and TNF-alpha. We added the estimated effect size in the manuscript and the table, as suggested, and explained that they weren't stated in the original studies but were estimated in our manuscript: "Two studies showed a statistically significant reduction: You et al. (2004) showed a reduction in CRP level of $p < 0.05$ (estimation 34% lower), with effect size estimated in 1.86 mg/L by subtracting the given values of before and after in table 2 and Ryan et al. (2014) showed a statistically significant reduction in CRP level of $p < 0.05$ (29% lower) showed and $p < 0.001$ (29% lower), with estimated effect size of 1.5 mg/L when comparing aerobic exercise and weight loss before and after graphs in figure 1 respectively. The effect size was estimated in both studies since the values weren't mentioned in the text or tables."*

- 3. Comment:** In the presentation of each parameter (in the Results/Outcomes: 1, 2, 3), it would be recommended to describe the methodology, i.e., which exercise methods were used, and then finish the paragraph with the results, i.e, how much increase or decreases were found.

Response: *Thank you for this feedback. We have updated the Result to state the methodology first and then the actual outcome.*

- 4. Comment:** The manuscript contains several spelling and punctuation issues. Also, American and British English were used interchangeably. It is recommended to use an automated tool to revise these issues in the manuscript before submitting.

Response: *We apologize for the spelling and punctuation errors. We have revised the final text in American English as the primary language using Grammarly and ChatGPT.*

Reviewer 3

- 1. Comment:** "Since the study aims to identify exercise as the main intervention, the use of specific terms like "aerobic exercise" or "resistance exercise" in the search strategy may have limited the number of documents that could have provided relevant information. In this regard, the manuscript title does not use "exercise" but "training" and includes terms such as "treadmill walking,"

"running," "jogging," etc. What if some studies mentioned these terms (and others included in the inclusion criteria) but did not specifically use the words "aerobic exercise" or "resistance exercise"? It is likely that several studies were not included. It would have been better to use "exercise" or "training" as search terms. Due to this limitation, it is recommended to review phrases like "exercise alone" in the "Key points" section because only studies that specifically refer to "aerobic exercise" or "resistance exercise" were included."

Response: Thank you for pointing this out. In addition to the use of specific terms "aerobic exercise" or "resistance exercise," we did include more search terms in our search strategy, as listed in detail in Appendix A. As an example for the search in Pubmed the following terms were included: "aerobic exercise"[All Fields] OR "aerobic training"[All Fields] OR "cardiovascular exercise"[All Fields] OR "endurance training"[All Fields] OR "Exercise"[MeSH Terms] OR training OR Exercise[All Fields] OR "physical activity"[All Fields] OR "Training"[All Fields] OR "resistance training"[All Fields] OR "resistance exercise"[All Fields]. Additionally, during the screening process, we specified which exercises or training were included as either resistance or aerobic interventions. For clarification, we have added a table summarizing all types of exercise/training with examples in Appendix E Table S1.

Appendix E

Table S1 Included and excluded types of interventions including synonymous terms and examples

Included types

Aerobic exercise

Synonyms include cardiovascular exercises, endurance training with aerobic component

Examples include tread mill, running, jogging, brisk walking, cycling, cardio classes, dancing, Zumba, rope jumping, hiking, swimming

Resistance exercise

Synonyms include strength training, circuit training with strength, bodyweight exercises

Examples include lifting weights, working with resistance bands, push-ups, squats, lunges, planks, pull-ups, jumping jacks, burpees.

Excluded types

flexibility exercise, e.g. stretching, yoga, pilates, thai chi AND/OR

balance exercise AND/OR

agility training, e.g. ladder drills, cone drills, jump training AND/OR

anaerobic exercise, e.g. sprinting, High-intensity training (HIIT) AND/OR

mixed modalities exercises/ training where aerobic and/or strength/resistance exercise is not analyzed separately

- 2. Comment:** "It is suggested to restructure the wording of all references to the study results to specifically relate them to the search terms used to specify the study intervention."

Response: We consistently used the terms "aerobic exercise" and "resistance exercise," even though different studies used other terms, such as "training," to refer to exercise per se, to keep them aligned with the search terms.

- 3. Comment:** "It is recommended not to use parentheses in the title."

Response: We thank the reviewer for this helpful suggestion. We have revised the title accordingly: "Effects of aerobic or resistance training on chronic inflammatory markers in postmenopausal women with overweight or obesity: A systematic review."

- 4. Comment:** "Please review the use of abbreviations in the abstract and the body of the document, ensuring that once the abbreviation is indicated, it is used consistently throughout the document."

Response: The text has been revised to ensure consistent use of abbreviations throughout the manuscript.

- 5. Comment:** “The third paragraph of the introduction is not referenced; please provide supporting information.”

Response: *Supporting citations have been added to the introduction after amendments were made.*

- 6. Comment:** “It is not clear why there is a citation at the end of the introduction, as this paragraph introduces the study objective, and the reason for this citation is unclear. Please clarify.”

Response: *We have moved to citations to the appropriate paragraph of the introduction.*

- 7. Comment:** “Please include the caption for Figure 1.”

Response: *We have revised the caption for Figure 1: “PRISMA Flow Diagram showing study identification, screening, eligibility assessment, and inclusion in the final review (n = 7)”.*

- 8. Comment:** “The results indicate that two non-randomized studies were included but non-controlled trials were excluded. What is the difference between these two types of studies?”

Response: *We included controlled trials with both an intervention and a control arm, encompassing randomized and non-randomized designs. Trials without a control group (non-controlled trials) were excluded because they do not allow direct comparison between the intervention and control conditions.*

- 9. Comment:** “Please start the discussion with your most important finding.”

Response: *Thank you for your comment. We have modified the discussion, starting with the most important findings.*

- 10. Comment:** Include among the limitations that your search restricted the study to an intervention only of "aerobic exercise" or "resistance exercise."

Response: *We have included this as a limitation of the study.*

Reviewer 4

- 1. Comment:** “Data synthesis is primarily descriptive and does not provide a deeper analytical interpretation. Tables reveal the results, although the narrative would benefit from a more clear description of the direction and magnitude of change, as well as whether any observed reductions are clinically meaningful. You should specifically state that exercise alone showed inconsistent results, and combined exercise with diet intervention proved effective. This difference is crucial for accurate interpretation.”

Response: *We have revised the Results section to highlight the different effects of exercise alone versus combined exercise and dietary interventions regarding each biomarker.*

- 2. Comment:** “The risk of bias should be managed more adequately. Although bias assessments have been made, the manuscript does not fully describe how high-risk studies influenced these findings. A sensitivity discussion or a comment on the weight for low-quality studies would enhance credibility.”

Response: *We have updated the Discussion to state that the results were interpreted cautiously, with reduced weight on studies assessed as having a high risk of bias. We have included the following sentence: “Given this risk-of-bias profile, the results were interpreted with caution, and findings from studies assessed as high risk were given less weight in forming the overall conclusions.”*

- 3. Comment:** “Moreover, randomized versus non-randomized trials also contributed to heterogeneity that limited their generalizability. This is something that the authors should clarify more explicitly, whether differences in study design contributed to the findings or not.”

Response: We have already updated the “Specific Differences Between Included Studies”, under “Difference in allocation strategy” sections with this information, as follows: “This systematic review consisted of five RCTs and two non-randomized trials. Regarding risk of bias, four RCTs and one non-RCT were rated as having some concerns, while one RCT and one non-RCT were rated as high risk of bias. Overall, the level of bias was comparable between study designs, and no major differences in outcomes were observed between the randomized and non-randomized trials.”

- 4. Comment:** “The variation in intervention design, sample size, and outcome measurement is significant. The combination of exercise and diet in the trials makes it difficult to isolate the effects of physical activity as an independent variable. The review would benefit from organizing results by intervention type, for example, exercise-only vs. combined, to facilitate a more precise interpretation.”

Response: We appreciate your insightful comment regarding the variations in intervention design and the importance of isolating the effects of physical activity. We understand the concern about trials combining exercise with other interventions. Our study selection process was specifically designed to address this. We only included trials where the effect of exercise could be clearly isolated. Even when additional interventions (such as diet or supplements) were present, we ensured the study design allowed for the independent assessment of the exercise component's impact. Therefore, due to the overall heterogeneity of each study's design regarding intervention type, we have kept the division of the tables regarding inflammatory markers.

- 5. Comment:** “Consistency among units of reporting, for example, pg/ml vs. pg/l , and means of reporting across tables is also needed to improve clarity.”

Response: Thank you for pointing these inconsistencies, we have fixed them throughout the text and tables for better clarity.

- 6. Comment:** The content is straightforward in theory, but some sections, especially the introduction and discussion, could be concisely rephrased without duplication. A more brief abstract-style summary of the main findings and limitations would also make it easier to read.

Response: We have updated the Introduction and Discussion sections as requested.

- 7. Comment:** Overall, this review is well-organized and poses a relevant question, but it requires refinement through in-depth analysis, explication, and convergence with the evidence presented.

Response: We sincerely appreciate the reviewer's constructive feedback. We have thoroughly revised the manuscript to incorporate more in-depth analysis, explication, and ensure stronger convergence with the evidence presented.

Reviewer 5

- 1. Comment:** “Title: The title is informative. I would do minor polishing, I suggest: “Effects of aerobic and/or resistance training on chronic inflammatory markers in postmenopausal women with overweight or obesity: A systematic review.”

Response: We thank the reviewer for this helpful suggestion. We have revised the title accordingly.

- 2. Comment:** “The introduction is clear, well-structured, there are minor spacing/punctuation issues need corrections.”

Response: We have revised the introduction accordingly and fixed these errors.

- 3. Comment:** “Search section: it’s a bit confusing. Right now it reads like you only searched literature from April 18 to May 1, 2025, which isn’t the case. The real meaning is:
- Search dates: You performed the searches between April 18 and May 1, 2025.
 - Coverage of literature: You included all studies from database inception through April 2025.
- Correct? If yes rewrite it.”

Response: *We thank the reviewer for this helpful suggestion. We have revised the Methods section to clarify the search dates and coverage, as follows: “A comprehensive literature search in PubMed, Web of Science, and the Cochrane Library was conducted between April 18 and May 1, 2025, covering all records from each database's inception through April 2025.”*

- 4. Comment:** “Aerobic and/or resistance exercise” — consider clarifying if “and/or” is appropriate throughout the manuscript.

Response: *We have carefully reviewed the manuscript and clarified instances of 'aerobic and/or resistance exercise,' retaining its use only where interventions genuinely encompassed both modalities or allowed for either, to accurately reflect the study designs.*

- 5. Comment:** There are sentences that are long; try to split them to enhance readability.

Response: *We have revised the text and split long sentences for better readability.*

- 6. Comment:** “Results: Avoid repeating exact numbers or mean ages that are already in the tables.”

Response: *Thank you. We have updated the results, including important numbers such as p-values when strictly necessary.*

- 7. Comment:** “Discussion: “This is the first systematic review regarding the impact of aerobic and resistance exercise on inflammatory markers in overweight and obese menopausal women.” However, you then reference Tan et al. (2023) and Khalafi et al. (2021), which already focus on postmenopausal overweight/obese women.”

Response: *We have removed the initial phrase that caused confusion and modified our Discussion section, mentioning both studies as follows: “Our findings contrast with those of Tan et al. (2023), whose meta-analysis of 34 studies (n = 2,229) concluded that exercise significantly reduced CRP, IL-6, and TNF- α levels. Their inclusion of short-term interventions (≥ 8 weeks) likely explains the discrepancy, as only eight of their studies lasted six months or longer—the minimum duration required in our review to capture long-term effects. Similarly, Khalafi et al. (2021) reviewed 32 studies (n = 1,510) and also found reductions in these biomarkers, but most interventions lasted only four to eight weeks.”*

- 8. Comment:** “Regarding risk of bias: Explain which domains were most problematic (e.g., randomization, allocation concealment, missing data) and how this could undermine confidence in study findings. Regarding the Language bias, highlight that restricting to English-language studies may exclude relevant trials from other regions, potentially skewing evidence and limiting generalizability.”

Response: *We thank the reviewer for this helpful suggestion. We have revised the discussion accordingly. We have addressed these in the following paragraphs:*

Under “Assessment of risk of bias in individual studies”: “The overall risk of bias profile supports a cautious interpretation of the results, particularly emphasizing the need to reduce the weight given to studies assessed at high or serious risk in any quantitative synthesis. The most frequently flagged domains across both assessment tools included concerns related to missing outcome data, selection of participants/randomization process, and deviations from intended interventions.” (...)

· Under “Strengths and Limitations”: “A comprehensive risk of bias assessment revealed substantial methodological concerns across most studies, compromising the validity of their results and reducing

confidence in our overall findings. Inherent research challenges, such as the impracticality of blinding participants in exercise groups, introduced potential for performance and detection bias. Given this risk-of-bias profile, the results were interpreted with caution, and findings from studies assessed as high risk were given less weight in forming the overall conclusions. Additionally, limiting the review to English-published articles may have excluded relevant trials from non-English-speaking regions, which could introduce language bias and reduce the generalizability of the findings.”

- 9. Comment:** “Regarding Heterogeneity: Elaborate on how differences in intervention type, intensity, duration, and participant inclusion criteria make comparisons across studies difficult and reduce the ability to generalize findings.”

Response: *This has been added to the results section, under the sub-heading “Specific Differences Between Included Studies,” as follows: “While our systematic review adhered to precise inclusion criteria, a detailed examination of the seven included studies revealed several sources of heterogeneity among their designs and participant characteristics, which warrant discussion as they influence the interpretation and generalizability of our findings.”*

- 10. Comment:** “Regarding Confounding factors: Explain that unmeasured or uncontrolled variables (diet, sleep, stress, alcohol) could influence inflammation independently of exercise, making it challenging to isolate exercise’s true effect.”

Response: *We thank the reviewer for this helpful suggestion. We have revised the discussion accordingly, by adding the following: “A critical limitation was the significant methodological and clinical heterogeneity across primary studies, particularly in intervention designs and comorbidity criteria, which complicated direct comparisons. Moreover, the interpretation of the effects of exercise was challenged by inconsistently assessed or controlled confounding factors such as broader lifestyle habits (e.g., diet, stress, sleep, alcohol), which could independently influence inflammation and make it difficult to isolate the true effect of exercise.”*

- 11. Comment:** “ The tables are informative and well-structured, but issues with unit consistency in table 4 and table 5 could confuse readers. Standardizing units, clarifying mean type, and simplifying footnotes would make it much clearer and journal-ready.”

Response: *We have addressed the suggestions and made the necessary edits to the tables.*

Thank you all once again for your valuable insights.