International Journal of Chronic Obstructive Pulmonary Disease

Dovepress
Taylor & Francis Group

Open Access Full Text Article

LETTER

Methodological Considerations Regarding Heterogeneity and Statistical Modeling in "Effects on Exercise Tolerance and Functional Outcomes of Eccentric versus Concentric Aerobic Exercise in People with COPD" [Letter]

Matías Otto-Yáñez[1], Rodrigo Torres-Castro[2], Jordi Vilaró[3]

[1]Health, Functionality, and Physical Activity Research Group (GISFAF), School of Kinesiology, Faculty of Health Sciences, Autonomous University of Chile, Santiago, Chile; [2]Department of Physical Therapy, University of Chile, Santiago, Chile; [3]Blanquerna School of Health Sciences, Global Research on Wellbeing, Ramon Llull University, Barcelona, Spain

Correspondence: Matías Otto-Yáñez, Email matias.otto@uautonoma.cl

Dear editor

We read with great interest the systematic review and meta-analysis by Magunagoikoetxea-Comins et al regarding the comparative effects of eccentric (ECC) and concentric (CON) aerobic exercise in patients with COPD.[1] While the study addresses a clinically relevant gap in pulmonary rehabilitation, we would like to offer some methodological observations that may influence the interpretation of the reported effect sizes and the certainty of the conclusions.

The authors state in their methodology that random-effects models would be employed to account for study heterogeneity. However, the forest plots are explicitly labelled as "IV, Fixed, 95% CI", indicating the use of fixed-effect inverse-variance models across outcomes. According to the Cochrane Handbook for Systematic Reviews of Interventions, the choice between fixed-effect and random-effects models is critical when heterogeneity is present.[2] When I^2 values exceed 75%, as seen in this study for Heart Rate ($I^2 = 93\%$), Dyspnea ($I^2 = 89\%$), and Endurance Time ($I^2 = 94\%$), the fixed-effect model tends to underestimate the standard error, leading to artificially narrow CIs and potentially inflated p-values. Applying a fixed-effect model under these conditions leads to "false precision," where a result appears highly significant despite substantial inconsistency across studies.[3]

To illustrate the potential impact of model choice, we contrasted the fixed-effect estimates reported by the authors with pooled effects re-estimated under random-effects assumptions (DerSimonian and Laird inverse-variance method) for each outcome using Review Manager (version 5.4, Cochrane Collaboration) software. For example, in Figure 3 of the article by Magunagoikoetxea-Comins et al, heart rate showed very high heterogeneity ($I^2 = 93\%$): while the fixed-effect model suggested a significant reduction (MD −14.37 bpm [−18.24, −10.50], $p<0.001$), the random-effects estimate was attenuated and became non-significant with a much wider CI (MD −11.14 [−26.26, 3.98], $p = 0.15$). Endurance time similarly presented extreme heterogeneity ($I^2=94\%$): the fixed-effect result was significant (SMD 1.05 [0.45, 1.64], $p<0.001$), but under random effects the CI widened substantially and crossed the null (SMD 1.06 [−1.54, 3.65], $p = 0.43$). Dyspnea (assessed by MRC) also had high heterogeneity ($I^2 = 89\%$); however, statistical significance was maintained under random effects (FE SMD −0.74 [−1.15, −0.34], $p<0.001$; RE SMD −1.44 [−2.74, −0.14], $p = 0.03$). Work rate exhibited the highest heterogeneity ($I^2 = 96\%$) and remained non-significant, with the random-effects estimate showing marked imprecision and even a change in direction (FE SMD −0.48 [−1.07, 0.12], $p=0.120$; RE SMD 2.89 [−1.41, 7.18], $p = 0.19$). Leg fatigue exhibited comparatively lower (though still notable) heterogeneity ($I^2 = 61\%$) and remained significant under both models (FE SMD −2.29 [−2.92, −1.67], $p<0.001$; RE SMD −2.59 [−3.71, −1.46], $p < 0.001$). Finally, 6MWT showed moderate heterogeneity ($I^2=66\%$) and remained non-significant under both approaches (FE SMD 0.17 [−0.30, 0.64], $p = 0.48$; RE SMD 0.24 [−0.58, 1.06], $p = 0.56$).

https://doi.org/10.2147/COPD.S590866
Received: 22 December 2025
Accepted: 10 March 2026
Published: 12 March 2026

As shown, for Endurance Time, the extreme I^2 of 94% and the proximity of the lower bound to zero suggest that a random-effects model would likely result in an interval crossing the null value, rendering the result non-significant. If random-effects models were applied as appropriate for outcomes with substantial heterogeneity ($I^2 \geq 60\%$), the GRADE assessment would require reconsideration. Under random effects, key outcomes that were significant under fixed effects—particularly heart rate ($I^2 = 93\%$) and endurance time ($I^2 = 94\%$)—become non-significant, with wider confidence intervals crossing the null, indicating increased imprecision per GRADE guidance. Combined with unexplained inconsistency due to very high heterogeneity, this would justify downgrading the certainty of evidence for these outcomes from moderate to low, yielding a more conservative and transparent interpretation of the findings.

Additionally, we noted the frequent use of the term "improvements" when describing outcomes that did not reach statistical significance (eg, quadriceps strength, $p = 0.43$, or 6MWT, $p = 0.12$). Standard reporting guidelines suggest that non-significant findings should be described as "no evidence of difference" to avoid biasing the reader toward a perceived benefit.[4] We also identified a data discrepancy in the abstract's conclusion, which attributes oxygen saturation benefits to ECC, while the data in Figure 3 of the article by Magunagoikoetxea-Comins et al clearly shows a significant advantage for the CON group ($p < 0.01$, $I^2 = 0\%$).[1]

We suggest that re-analyzing outcomes with high heterogeneity using a random-effects model, alongside a cautious application of the GRADE framework,[5] could offer a more conservative and nuanced perspective on the effects of eccentric exercise. Such an approach may enhance the manuscript's contribution to the field by providing clinicians with a more tempered expectation of the functional outcomes associated with these exercise modalities in people with COPD.

Abbreviations

COPD, chronic obstructive pulmonary disease; ECC, eccentric aerobic exercise; CON, concentric aerobic exercise; MD, mean difference; CI, confidence interval; TUG, timed up and go; 6MWT, six-minute walk test; GRADE, Grading of Recommendations Assessment, Development and Evaluation.

Disclosure

The authors report no conflicts of interest in this communication.

References

1. Magunagoikoetxea-Comins G, Jiménez-García ME, Pérez-Ferreiro M, Fernandez-Pardo TE. Effects on exercise tolerance and functional outcomes of eccentric versus concentric aerobic exercise in people with COPD: a systematic review and meta-analysis. *Int J Chron Obstruct Pulmon Dis*. 2025;20:4061–4078. doi:10.2147/COPD.S558167
2. Deeks JJ, Higgins JPT, Altman DG. Chapter 10: analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd ed. John Wiley & Sons; 2019. doi:10.1002/9781119536604.ch10
3. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to Meta-Analysis*. John Wiley & Sons; 2021. doi:10.1002/9780470743386
4. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539–1558. doi:10.1002/sim.1186
5. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924–926. doi:10.1136/bmj.39489.470347.AD

https://doi.org/10.2147/COPD.S590866