# Heterogenous treatment effects following inspiratory muscle training during recovery from post-acute COVID-19 syndrome.

METCALFE, R.S., SWINTON, P.A., MACKINTOSH, K.A., BERG, R.M.G., SHELLEY, J., SAYNOR, Z.L., HUDSON, J., DUCKERS, J., LEWIS, K., DAVIES, G.A. and MCNARRY, M.A.

2023

# Heterogenous Treatment Effects Following Inspiratory Muscle Training during Recovery from Post-Acute COVID-19 Syndrome

Richard S Metcalfe[1], Paul A Swinton[2], Kelly A Mackintosh[1], Ronan M. G. Berg[3-6], James Shelley[1,7], Zoe L Saynor[8], Joanne Hudson[1], Jamie Duckers[9], Keir Lewis[10,11], Gwyneth A Davies[11,12], Melitta A McNarry[1,*]

[1] Applied Sports, Technology, Exercise and Medicine (A-STEM) Research Centre, Swansea University, Swansea, UK

[2] School of Health Sciences, Robert Gordon University, Aberdeen, UK

[3] Department of Biomedical Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[4] Centre for Physical Activity Research, University Hospital Copenhagen – Rigshospitalet, Copenhagen, Denmark

[5] Department of Clinical Physiology and Nuclear Medicine, University Hospital Copenhagen – Rigshospitalet, Copenhagen, Denmark

[6] Neurovascular Research Laboratory, Faculty of Life Sciences and Education, University of South Wales, Pontypridd, UK

[7] Lancaster Medical School, Lancaster University, Lancaster, UK

[8] Physical Activity, Health and Rehabilitation Thematic Research Group Lead, School of Sport, Health and Exercise Science, Faculty of Science and Health, University of Portsmouth, Portsmouth, UK.

[9] All Wales Adult CF Centre, Cardiff and Vale University Health Board, Cardiff, UK

[10] Hywel Dda University Health Board, UK.

[11] School of Medicine, Swansea University, Swansea, UK

[12] Swansea Bay University Health Board, UK.

* Correspondence to: Professor Melitta McNarry. Applied Sports, Technology, Exercise and Medicine Research Centre, Swansea University, Swansea, UK. Email: m.mcnarry@swansea.ac.uk.

**ABSTRACT**

**Purpose:** To investigate whether heterogeneous treatment effects occur for changes in inspiratory muscle strength, perceived dyspnoea, and health-related quality of life (QoL), following eight-weeks unsupervised home-based inspiratory muscle training (IMT) in adults with post-acute COVID-19 syndrome. **Methods**: In total, 147 adults with self-reported prior COVID-19 either completed an eight-week home-based IMT intervention (*n*=111; 92 females; 48±11 years; 9.3±3.6 months post-acute COVID-19 infection) or acted as "usual care" wait list controls (*n*=36; 34 females; 49±12 years; 9.4±3.2 months post-acute COVID-19 infection). **Results:** Applying a Bayesian framework, we found clear evidence of heterogeneity of treatment response for inspiratory muscle strength: the estimated difference between standard deviations (SDs) of the IMT and control groups was 22.8 $cmH_2O$ (75% Credible Interval (CrI): 4.7-37.7) for changes in maximal inspiratory pressure (MIP), and 86.8 pressure time-units (PTUs; 75% CrI: 55.7-116.7) for sustained MIP (SMIP). Conversely, there were minimal differences in the SDs between the IMT and the control group for changes in perceived dyspnoea and health-related QoL, providing no evidence of heterogeneous treatment effects. Higher cumulative power during the IMT intervention was related to changes in MIP ($\beta$=10.9 [95% CrI: 5.3-16.8] $cmH_2O$ per 1SD) and SMIP ($\beta$=63.7 [32.2-95.3] PTUs per 1SD), clearly indicating an IMT dose response for changes in inspiratory muscle strength. Older age (>50 years), a longer time post-acute COVID-19 (>3 months), and greater severity of dyspnoea at baseline were also associated with smaller improvements in inspiratory muscle strength. **Conclusion:** Heterogenous individual responses occurred following an eight-week home-based IMT programme in people with post-acute COVID-19 syndrome. Consistent with standard exercise theory, larger improvements in inspiratory muscle strength are strongly related to a greater cumulative dose of IMT.

**Keywords:** post-acute COVID-19 syndrome, long COVID, rehabilitation, treatment, breathlessness, breathing

## INTRODUCTION

Post-acute Coronavirus Disease19 (COVID-19) syndrome (1), often referred to as long COVID, is estimated to affect 1 in 10 individuals with COVID-19, which in the UK, equates to ~2.3 million people as of October 6th 2022 (2). Whilst the symptoms of post-acute COVID-19 syndrome are diverse and vary between individuals, breathlessness is amongst the most common and debilitating (3). Given the prevalence and burden of post-acute COVID-19 syndrome, there is a need to develop feasible and effective rehabilitation strategies, emphasised by recent evidence that prior vaccination only partially protects against developing post-acute COVID-19 syndrome (4).

There are currently limited rehabilitation strategies available for people with post-acute COVID-19 syndrome, but a recent randomised controlled trial demonstrated that inspiratory muscle training (IMT) is an effective intervention to enhance recovery from COVID-19 (5). IMT involves repeated inspiratory breaths performed using a resisted air-flow device and is designed to challenge and elicit adaptations in the respiratory musculature (6). Following eight weeks of unsupervised home-based IMT, there were mean improvements in perceived dyspnoea, inspiratory muscle strength, device-measured moderate-intensity physical activity, and estimated aerobic fitness (5). IMT is low cost and simple to deliver remotely, making it ideal to integrate as part of a multi-component rehabilitation programme for people with post-acute COVID-19 syndrome.

The efficacy of a treatment, including exercise interventions, is typically presented as a mean change compared to a control group, yet this approach may overlook potentially important individual differences in the response to the intervention, which are referred to as ‘heterogenous treatment effects’ (7). Quantifying and predicting such inter-individual variation is the basis of precision medicine, which aims to prescribe individually tailored

interventions to optimise treatment outcomes (8,9). Nevertheless, whether meaningful individual variation in response to either supervised or unsupervised exercise training truly exists is somewhat contentious; many previous studies have not applied statistical frameworks that account for technical, biological, and random error (10,11). Specifically, to be able to conclude that true individual differences in the response to the intervention exist, there must be evidence of larger variation in the change scores in the intervention group compared to the change scores from an appropriate time-matched control group (10). If this is the case, it would be appropriate to subsequently explore moderating factors that may explain the additional variation in response in the intervention group (10). Conversely, if the variation of change is similar between the intervention and the control groups, then it is not possible to conclude that there were any individual responses caused by the intervention *per se* (10). Whilst there may still be a mean intervention effect, and large variation around the mean change in the intervention group, it could only be concluded that this was caused by factors present in both the intervention and control groups (i.e., technical, biological, or random error; (10)).

Any variation in treatment effects may be more likely to be present and/or be more pronounced in studies of the real-world effectiveness of interventions, particularly involving home-based exercise, where the lack of supervision could result in large differences in individual adherence to the prescribed intervention (i.e. intervention fidelity) (12). Furthermore, the diverse presentation of post-acute COVID-19 syndrome (3,13) may lead to large inter-individual treatment effects. Thus, the primary aim of this study was to investigate whether heterogeneous treatment effects occur following eight weeks of unsupervised IMT in adults with post-acute COVID-19 syndrome. Where heterogenous treatment effects were identified, two secondary aims were to: *1)* quantify the proportion of individuals expected to make an improvement following IMT; and *2)* perform sub-analyses on participant

50 characteristics and IMT dose-related variables to explore relative treatment effect
51 modification.

## METHODS

### *Participants*

The sample for this study is from an eight-week, single-centre, two-arm randomised controlled trial (RCT) which investigated the effect of home-based IMT on inspiratory muscle function, self-reported health status, and physical activity levels, in adults with self-reported post-acute COVID-19 syndrome (5). The mean intervention effects are presented in McNarry *et al.* (5). For this secondary analysis of potential heterogeneous treatment effects, of 281 participants originally randomised, we excluded all participants who did not complete the study and/or had incomplete outcome data ($n$=134). This resulted in a sample of 147 participants who were randomised to either the IMT ($n$=111) or a "usual care" wait list control ($n$=36) group (**Table 1**). All participants provided informed consent following approval by the NHS Research Ethics Committees (Ref: 20/HRA/3536). The study was pre-registered on the Health and Care Research Wales Research Directory (Ref: 48075) and conducted in accordance with the Declaration of Helsinki.

### *Outcomes*

Inspiratory muscle strength was measured at baseline and post-intervention using a handheld inspiratory resistive flow device ($PrO_2^{TM}$, $PrO_2$Fit Health Incorporated, RI, USA). Following familiarisation with the device, participants performed full expiration to residual volume, followed by a maximal sustained inspiratory effort to measure both maximal inspiratory pressure (MIP) and sustained maximal inspiratory pressure (SMIP). The assessment was performed at home and supervised via remote teleconference (due to lockdown) with strong verbal encouragement provided. Both MIP and SMIP are important clinical markers of respiratory function (14), which the $PrO_2^{TM}$ device measures with high reliability (15).

Changes in dyspnoea were assessed using the Transition Dyspnoea Index (TDI; (16)) and the 15-item King's Brief Interstitial Lung Disease (KBILD) Questionnaire (17). The TDI is a clinically validated questionnaire, which measures changes in dyspnoea from baseline using the Baseline Dyspnoea Index (BDI) in three domains (functional impairment, magnitude of task, and magnitude of effort), and was completed post-intervention only. The KBILD was completed at baseline and post-intervention and provides a score for overall health-related quality of life from responses within three sub-categories (Psychological, Breathlessness and Activities, and Chest Symptoms).

***IMT Intervention***

Participants randomised to the IMT group were prescribed an eight-week home-based IMT intervention, with a frequency of three sessions per week performed on non-consecutive days. IMT training was delivered using the same $PrO_2^{TM}$ device that was used for assessing MIP and SMIP. Training in the use of the $PrO_2^{TM}$ device was provided to each participant during a one-to-one video conferencing meeting.

Each IMT session lasted ~20 minutes. Participants were prescribed a maximum of six blocks of six inspirations, with each breath interspersed with a short period of resting recovery which progressively decreased from 40 seconds to 10 seconds within each distinct block. Each inspiratory breath was performed at >80% of SMIP ascertained from a maximal inspiratory effort, performed prior to each IMT session to allow for both training progression, as well as potential day-to-day fluctuations in respiratory function due to the relapsing/remitting nature of post-acute COVID-19 syndrome (3). Each inspiration was performed for as long as possible and, during each IMT session, participants completed as many inspirations as they could prior to failure, defined as not achieving 80% SMIP on three consecutive breaths.

The $PrO_2^{TM}$ device synchronises wirelessly to a computer, smartphone or tablet via an application (https://apps.apple.com/us/app/pro2-fit/id1321623265), which provided real-time graphic biofeedback during each session. This also facilitated remote recording and cloud storage of the characteristics of all participants' training sessions. The following characteristics of the IMT training sessions were subsequently extracted: *(1)* total completed sessions; *(2)* mean training frequency (sessions/week over eight weeks); *(3)* total breaths; *(4)* mean number of breaths per session; *(5)* mean breath duration; and *(6)* total cumulative power across the intervention. As total cumulative power across the intervention could be influenced by baseline MIP and/or SMIP, this was expressed in both absolute terms and relative to baseline MIP and SMIP in the analysis. The variation in IMT training characteristics is shown in **Table 2.**

***Statistical Analysis***

All analyses were conducted within a Bayesian framework and are reported in accordance with the CHAMP statement (18). Seven dependent variables were selected, including the TDI, KBILD and its sub-categories, and MIP and SMIP. Individual change scores were calculated by subtracting baseline from post-intervention values (except for the TDI where the post-intervention score reflects change from baseline). Variation in change scores were compared across the intervention and control group with greater standard deviations for the intervention group taken as evidence of heterogeneous treatment effects. Distributional models estimating mean and variance parameters were fitted for each dependent variable, either including group as a predictor for the standard deviation ($M_2$), or not ($M_1$). Bayes factors $\left(\frac{p(y|M_1)}{p(y|M_2)}\right)$ were calculated with the strength of evidence in favour of $M_1$ (no heterogeneous treatment effects) or $M_2$ (heterogeneous treatment effects) assessed according to a previously defined scale (19). The data-generating model for each variable was assessed

by fitting normal, skew normal, and *t*-distributions with the most appropriate distribution type for each outcome determined using the Watanabe-Akaike information criterion. Model checking and selection was performed to increase the precision of results. Differences in standard deviation between the intervention and control were estimated using posterior predictions and 95% credible intervals (CrIs).

Where strong evidence of heterogenous treatment effects was obtained (Bayes factor >10), proportion of response and factors associated with relative treatment effect modification were explored. Proportion of response was estimated by subtracting the mean difference between groups and the difference in the standard deviations to calculate the intervention-response standard deviation and calculating the proportion of the distribution exceeding zero (20). Subgroups comprising binary classification of patient characteristics (time since COVID [low: ≤ 3 months; high: >3 months], Body Mass Index (BMI) [low: <25 $kg•m^{-2}$; high: ≥25 $kg•m^{-2}$], age [low: <50 years; high: ≥50 years], baseline KBILD total score [low: <53; high: ≥53], and baseline BDI [low: ≤6 units; high: >6 units]) were created and the difference in mean treatment-effect estimated, with Bayes factors and 95% CrIs calculated to interpret relative treatment-effect modification. For age, baseline KBILD and baseline BDI score, low and high scores were split based on the median, whilst for BMI the standard overweight cut-off of 25 $kg•m^{-2}$ was applied and for time since COVID a 3-month cut-off was applied based on the World Health Organisation clinical case definition of long COVID (21). For IMT-related variables, relative treatment effect modification was assessed by linearly regressing change scores on each variable standardised by dividing by the sample standard deviation. Default weakly informative Student-*t* prior and half-*t* priors with three degrees of freedom were used for intercept and variance parameters (22). All analyses were performed using the R wrapper package brms interfaced with Stan to perform sampling (23) and the R package

147 bridge sampling to calculate Bayes factors. Convergence of parameter estimates was obtained
148 for all models with Gelman-Rubin R-hat values below 1.1 (24).

## Results

### *KBILD and TDI*

The best model fit for KBILD sub-domain and total score, and the TDI score, was obtained using a normal distribution. There were minimal differences in the standard deviation scores between the IMT and the control group for all KBILD sub-domains (**Figure 1**), the KBILD total score (**Figure 1**), and the TDI score (**Figure 2**), and in all cases the Bayes factor was <3, providing no evidence of individual responses to IMT (**Table 3**).

### *MIP and SMIP*

The best model fit for MIP was a *t*-distribution and for SMIP it was a normal distribution. The estimated difference in standard deviations of the IMT and the control group was 22.8 $cmH_2O$ (75% CrI: 4.7-37.7) for MIP, and 86.8 pressure time-units (75% CrI: 55.7-116.7) for SMIP. In both cases, the Bayes factor was >100, providing extreme evidence of individual responses to IMT (**Figure 3**, **Table 3**). The estimated proportion of response was 0.84 (95% CrI: 0.63-1.0) for MIP and 0.95 (95% CrI: 0.76-1.0) for SMIP (**Table 3**).

There was evidence of an IMT dose-response: a greater treatment effect for both MIP and SMIP was shown with a higher number of IMT sessions (moderate evidence for both), more breaths performed per session (extreme evidence for MIP, very strong evidence for SMIP), a larger number of total breaths performed over the intervention (very strong evidence for both), a higher mean breath duration (strong evidence for MIP, very strong evidence for SMIP), a higher total cumulative power expressed absolutely or relative to baseline MIP/SMIP (all extreme evidence) (**Table 4**).

Several participant characteristics also appeared to alter the treatment effect: the change in both MIP and SMIP was greater in younger participants (extreme evidence for MIP, strong

evidence for SMIP), those with COVID-19 less than 3 months before baseline assessment (strong evidence for both), and those who had less severe dyspnoea at baseline (extreme evidence for MIP, strong evidence for SMIP; **Figure 4**). There was less evidence that treatment effect was altered by baseline BMI (anecdotal evidence favouring no effect for MIP, moderate evidence for an effect for SMIP) or baseline KBILD total score (moderate evidence for both; **Figure 4**). Point estimates from standard Pearson correlations identified limited, but likely non-zero associations between age, baseline dyspnoea (BDI score) or time since COVID-19 infection and any IMT dose variable ($r$<0.3).

## DISCUSSION

This study investigated whether heterogeneity of treatment effects occurs following eight weeks of unsupervised home-based IMT in adults recovering from COVID-19. There were three key findings: *1)* there were clear heterogeneous treatment effects for changes in respiratory muscle strength, and consistent with standard exercise theory, larger improvements were related to a greater accumulated dose of IMT (more sessions, more breaths, greater cumulative power etc.); *2)* improvements in respiratory muscle strength following IMT were lower in participants who were older, when IMT was initiated >3 months following onset of COVID-19, and in participants with more severe dyspnoea at baseline; and *3)* for changes in perceived dyspnoea and health-related quality of life, there was large between-participant variability in both the IMT and control groups, but no evidence of heterogeneous IMT treatment effects.

This is the first investigation of heterogeneity in the *effectiveness* of an unsupervised, home-based exercise or physical activity intervention using appropriate statistical methods. Whilst several recent studies have been unable to detect heterogeneous treatment effects for body composition, cardiorespiratory fitness and blood pressure following *supervised* exercise training in adults (25–28), we found extremely strong evidence of individual responses for changes in inspiratory muscle strength following unsupervised IMT. This discrepancy is likely explained, at least in part, by the additional variability in intervention adherence and fidelity present in our study given its remote and unsupervised delivery method. Indeed, the improvements in inspiratory muscle strength were positively related to IMT characteristics, including number of training sessions, number of training breaths, the duration of training breaths, and total cumulative power over the intervention. Total cumulative power was the strongest predictor of changes in inspiratory muscle strength: for every 1 SD increase in total

cumulative power over the eight-week intervention, we observed a further improvement in MIP of 10.9 [95% CrI: 5.3-16.8] cm $H_2O$ and a further improvement in SMIP of 63.7 [95% CrI: 32.2-95.3] PTUs. These data provide the clearest evidence of a dose-response relationship for improvements in inspiratory muscle strength following IMT. This finding can inform the delivery of IMT as a rehabilitative tool for post-acute COVID-19 syndrome, but it is likely that similar heterogeneity would also be observed with the delivery of home-based IMT in other chronic respiratory conditions where IMT has been shown to be beneficial for inspiratory muscle strength, such as chronic obstructive pulmonary disease (29), asthma (30) or cystic fibrosis (31). There are a wide range of individual, psychosocial, and disease-specific factors that influence adherence to home-based exercise (e.g., 32,33) and it will be important for future research to determine the potential barriers and facilitators that influence adherence to unsupervised home-based exercise in people with post-acute COVID-19 syndrome.

In comparison to previous studies of heterogeneity of exercise response, this study investigated a population living with a disease of highly diverse manifestation and aetiology (3,13). Our findings suggest this impacted the improvements in inspiratory muscle strength following IMT. Specifically, we found that more severe dyspnoea and initiating IMT >3 months following COVID-19 infection, were related to smaller improvements in inspiratory muscle strength. Such findings have implications for the timing of rehabilitation components, suggesting that IMT should be offered early in rehabilitation programmes to maximise its efficacy. Interestingly, whilst we observed a dose-response to IMT within the sample as a whole, there were no notable correlations between baseline dyspnoea (BDI score), time since COVID-19 infection, or age, and any IMT dose variable. This implies that the smaller improvements were due to differences in the physiological response to a given dose of IMT

in these subpopulations, rather than systematic differences in the quantity/quality of IMT exposure.

It is curious that we could detect evidence of meaningful heterogeneity of response for physiological outcomes (respiratory muscle strength) but not for subjective outcomes (perceived dyspnoea and health-related quality of life). It is, however, noteworthy that there was a high level of between-participant variability for these subjective outcomes in both the control and intervention group. This is perhaps unsurprising given the relapsing/remitting nature of post-acute COVID-19 syndrome symptomatology (3), together with evidence that subjective measures of dyspnoea can be unrelated to underlying disease severity and influenced by multiple other situational factors (e.g., emotional, behavioural, environmental; (34)). As such, the lack of observed heterogeneity in treatment response for these outcomes may be partly explained by the typical error of measurement generally being higher for the subjective, compared to the physiological, outcomes in our study. The typical error (20) expressed relative to the baseline standard deviation, was 0.45 for MIP, 0.35 for SMIP, and 0.54, 0.44, 0.76 and 0.62 for the KBILD breathlessness, psychological, chest, and total scores, respectively. Therefore, we cannot specifically rule out meaningful heterogenous treatment effects for these subjective outcomes, but high measurement error will inevitably mask any individual variability in the treatment group and make identification of potential moderator/mediator variables challenging. It is also important to note that there was a mean improvement in perceived dyspnoea (TDI score) with IMT compared to the control group (5). Thus, if prescribing IMT as a rehabilitative intervention, it would be prudent to aim to maximise improvements in clinically relevant physiological outcomes in the knowledge that some improvement in perceived dyspnoea will also likely be exhibited.

**Practical Implications**

MIP and SMIP as markers of respiratory muscle function are recognised as important clinical outcomes in people with pulmonary disease (14). IMT could be part of a therapeutic programme for people with pulmonary disease (29,30), including post-acute COVID-19 syndrome (5), and our findings demonstrate that prescription of unsupervised home-based IMT - the likely scenario for scalable real-world implementation - leads to heterogeneous responses for changes in respiratory muscle strength. A high proportion of people can be expected to see some change in respiratory muscle strength following IMT (~84% for MIP and ~95% for SMIP), but greater improvements are observed with a larger dose of IMT. Therefore, practitioners who are implementing IMT as a rehabilitation tool in people with post-acute COVID-19 syndrome should encourage patients to accumulate a larger dose of training to maximise improvements in inspiratory muscle strength. Our findings also provide the basis for future research to determine: 1) why older age, a longer time post-acute COVID-19 syndrome, and a greater severity of baseline dyspnoea, are associated with smaller improvements in inspiratory muscle strength following IMT; and 2) how IMT may be adapted to enhance the improvements in these subpopulations.

**Limitations**

Whilst there are numerous strengths of this study, certain limitations need to be acknowledged. Firstly, we applied conventional subgroup analysis to identify potential moderators of individual responses for respiratory muscle strength (7). Although this approach can identify theoretical conditions under which the intervention is most/least effective, there are limitations to its use to inform decision making at an individual level; individuals can belong to multiple different subgroups which may yield different inferences about the optimal treatment effect (7). There are also potential relationships between the mean and standard deviation of change scores (35), such that some of the apparent

heterogenous treatment effects in inspiratory muscle strength may reflect systematic changes in the intervention versus control group. Our population was also largely female and, although this is reflective of a higher female prevalence of post-acute COVID-19 syndrome (36), it was not possible to determine whether heterogeneity in response would be present in males, or whether biological sex is a potential moderator of the heterogeneity. It should also be noted that the questionnaires utilised in this study have not specifically been validated in people with post-acute COVID-19 syndrome. In addition, we took the decision to focus on dyspnoea and have not collected data on the range or severity of other symptoms that were experienced (3). Finally, as the data was collected entirely remotely and during periods of lockdown, there were limitations on the outcome measures able to be obtained. Whilst changes in MIP and SMIP are key markers of pulmonary function, it will be important to determine whether similar heterogeneity is present for other markers (e.g. diaphragm thickness, ventilatory reserve etc.).

## Conclusions

We have previously reported that eight weeks of unsupervised home-based IMT resulted in *mean* improvements in perceived dyspnoea and inspiratory muscle strength in people with post-acute COVID-19 syndrome (5). The present findings provide additional novel insight by demonstrating that there is *individual variability* in the improvement in inspiratory muscle strength (but not perceived dyspnoea) following IMT in people recovering from COVID-19 (i.e., some people get more benefit, and some people get less benefit from the IMT intervention for inspiratory muscle strength). Consistent with standard exercise theory, larger improvements in clinically relevant markers of inspiratory muscle strength are strongly related to a greater cumulative dose of IMT over the intervention.

## AUTHOR CONTRIBUTIONS

MAM and KAM conceived the idea for the primary RCT and were the grant holders and principal investigators. MAM, KAM, ZLS, GAD, KL, JD, RB, JH were involved in the design of the primary RCT. MM and JS collected the data for the primary RCT. RSM, MAM, KAM and PAS conceived the idea for this manuscript. PAS provided statistical expertise and performed the statistical analysis. RM wrote the initial draft of the manuscript. All authors were involved in drafting versions and critically revising for important intellectual content. All authors have read and approved the final version. MAM is the guarantor of the study.

## ACKNOWLEDGEMENTS

We would like to thank the participants who took part in this research. The results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation, and statement that results of the present study do not constitute endorsement by ACSM.

## FUNDING

This research was funded by the Welsh Government Sêr Cymru III Tackling COVID-19 grant scheme (Reference MA/KW/1457/20) and The Higher Education Funding Council for Wales Research Wales Innovation Fund (Collaboration Booster Faculty Fund), grant number FF4. The Centre for Physical Activity Research is supported by TrygFonden grants (ID 101390, ID 20045 and ID 125132).

## COMPETING INTERESTS

None to declare.

## DATA AVAILABILITY STATEMENT

The deidentified data are available from the corresponding author upon reasonable request.

## References


1. Nalbandian A, Sehgal K, Gupta A, et al. Post-acute COVID-19 syndrome. *Nat Med* 2021; 27(4):601–15. https://doi.org/10.1038/s41591-021-01283-z

2. Prevalence of ongoing symptoms following coronavirus (COVID-19) infection in the UK. *Office for National Statistics*. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/prevalenceofongoingsymptomsfollowingcoronaviruscovid19infectionintheuk/6october2022

3. Davis HE, Assaf GS, McCorkell L, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *eClinicalMedicine* 2021; 38: 101019. https://doi.org/10.1016/j.eclinm.2021.101019

4. Al-Aly Z, Bowe B, Xie Y. Long COVID after breakthrough SARS-CoV-2 infection. *Nat Med* 2022; 25: 1–7. https://doi.org/10.1038/s41591-022-01840-0

5. McNarry MA, Berg RMG, Shelley J, et al. Inspiratory Muscle Training Enhances Recovery Post COVID-19: A Randomised Controlled Trial. *Eur Respir Journal* 2022; 60: 2103101. https://doi.org/10.1183/13993003.03101-2021

6. Enright SJ, Unnithan VB, Heward C, Withnall L, Davies DH. Effect of High-Intensity Inspiratory Muscle Training on Lung Volumes, Diaphragm Thickness, and Exercise Capacity in Subjects Who Are Healthy. *Phys Ther* 2006; 86(3):345–54. https://doi.org/10.1093/ptj/86.3.345

7. Kent DM, Steyerberg E, Klaveren D van. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. *BMJ*. 2018; 10; 363:k4245. https://doi.org/10.1136/bmj.k4245

8. Collins FS, Varmus H. A New Initiative on Precision Medicine. *N Engl J Med* 2015; 372(9):793–5. https://doi.org/10.1056/NEJMp1500523

9. Denny JC, Collins FS. Precision medicine in 2030—seven ways to transform healthcare. *Cell* 2021; 184(6):1415–9. https://doi.org/10.1016/j.cell.2021.01.015

10. Atkinson G, Batterham AM. True and false interindividual differences in the physiological response to an intervention. *Exp Physiol* 2015; 100(6):577–88. https://doi.org/10.1113/EP085070

11. Williamson PJ, Atkinson G, Batterham AM. Inter-Individual Responses of Maximal Oxygen Uptake to Exercise Training: A Critical Review. *Sports Med* 2017; 47(8):1501–13.

12. Nicolson PJA, Hinman RS, Kasza J, Bennell KL. Trajectories of adherence to home-based exercise programs among people with knee osteoarthritis. *Osteoarthritis Cartilage* 2018; 26(4):513–21. https://doi.org/10.1016/j.joca.2018.01.009

13. Tabacof L, Tosto-Mancuso J, Wood J, et al. Post-acute COVID-19 Syndrome Negatively Impacts Physical Function, Cognitive Function, Health-Related Quality of Life, and Participation. *Am J Phys Med Rehabil* 2022; 101(1):48–52. https://doi.org/10.1097/PHM.0000000000001910

14. Formiga MF, Vital I, Urdaneta G, Campos MA, Cahalin LP. Beyond inspiratory muscle strength: Clinical utility of single-breath work capacity assessment in veterans with COPD. *Respir Med* 2019; 147:13–8. https://doi.org/10.1016/j.rmed.2018.12.012

15. Formiga MF, Roach KE, Vital I, et al. Reliability and validity of the test of incremental respiratory endurance measures of inspiratory muscle performance in COPD. *Int J Chron Obstruct Pulmon Dis* 2018; 13:1569–76. https://doi.org/10.2147/COPD.S160512

16. Mahler DA, Tomlinson D, Olmstead EM, Tosteson AN, O'Connor GT. Changes in dyspnea, health status, and lung function in chronic airway disease. *Am J Respir Crit Care Med* 1995; 151(1):61–5. https://doi.org/10.1164/ajrccm.151.1.7812573

17. Patel AS, Siegert RJ, Brignall K, et al. The development and validation of the King's Brief Interstitial Lung Disease (K-BILD) health status questionnaire. *Thorax* 2012; 67(9):804–10. https://doi.org 10.1136/thoraxjnl-2012-201581

18. Mansournia MA, Collins GS, Nielsen RO, et al. A CHecklist for statistical Assessment of Medical Papers (the CHAMP statement): explanation and elaboration. *Br J Sports Med* 2021; 55(18):1009–17. http://dx.doi.org/10.1136/bjsports-2020-103652

19. Jeffreys H. The Theory of Probability. OUP Oxford; 1939.

20. Swinton PA, Hemingway BS, Saunders B, Gualano B, Dolan E. A Statistical Framework to Interpret Individual Response to Intervention: Paving the Way for Personalized Nutrition and Exercise Prescription. *Front Nutr* 2018; 5:41. https://doi.org/10.3389/fnut.2018.00041

21. Soriano JB, Murthy S, Marshall JC, Relan P, Diaz JV. A clinical case definition of post-COVID-19 condition by a Delphi consensus. *Lancet Infect Dis* 2022; 22(4):e102–7. https://doi.org/10.1016/S1473-3099(21)00703-9

22. Gelman A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal* 2006; 1(3):515–34. https://doi.org/10.1214/06-BA117A

23. Bürkner PC. brms: An R Package for Bayesian Multilevel Models Using Stan. *J Stat Softw*. 2017; 80(1): 1-28. https://doi.org/10.18637/jss.v080.i01

24. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis Third edition (with errors fixed as of 15 February 2021). Chapman and Hall/CRC. https://doi.org/10.1201/b16018

401 25. Williamson PJ, Atkinson G, Batterham AM. Inter-individual differences in weight
402 change following exercise interventions: a systematic review and meta-analysis of
403 randomized controlled trials. *Obes Rev* 2018; 19(7):960–75.
404 https://doi.org/10.1111/obr.12682

405 26. Kelley GA, Kelley KS, Pate RR. Are There Inter-Individual Differences in Fat Mass
406 and Percent Body Fat as a Result of Aerobic Exercise Training in Overweight and
407 Obese Children and Adolescents? A Meta-Analytic Perspective. *Child Obes* 2020;
408 16(5):301–6. https://doi.org/10.1089/chi.2020.0056

409 27. Kelley GA, Kelley KS, Stauffer BL. Walking and resting blood pressure: An inter-
410 individual response difference meta-analysis of randomized controlled trials. *Sci Prog*
411 2022;105(2). https://doi.org/10.1177/00368504221101636

412 28. Bonafiglia JT, Swinton PA, Ross R, et al. Interindividual Differences in Trainability and
413 Moderators of Cardiorespiratory Fitness, Waist Circumference, and Body Mass
414 Responses: A Large-Scale Individual Participant Data Meta-analysis. *Sports Med* 2022.
415 https://doi.org/10.1007/s40279-022-01725-9

416 29. Gosselink R, Vos JD, Heuvel SP van den, Segers J, Decramer M, Kwakkel G. Impact of
417 inspiratory muscle training in patients with COPD: what is the evidence? *Eur Respir J*
418 2011;37(2):416–25. https://doi.org/10.1183/09031936.00031810.

419 30. Silva IS, Fregonezi GAF, Dias FAL, Ribeiro CTD, Guerra RO, Ferreira GMH.
420 Inspiratory muscle training for asthma. *Cochrane Database Syst Rev* 2013;
421 9:CD003792. https://doi.org/10.1002/14651858.CD003792.pub2

422 31. Stanford G, Ryan H, Solis-Moya A. Respiratory muscle training for cystic fibrosis.
423 *Cochrane Database Syst Rev* 2020; 12:CD006112.
424 https://doi.org/10.1002/14651858.CD006112.pub5

425 32. Hancox, JE, van der Wardt V, Pollock K, Booth V, Vedhara K, Harwood RH. Factors
426 influencing adherence to home-based strength and balance exercises among older adults
427 with mild cognitive impairment and early dementia: Promoting Activity, Independence
428 and Stability in Early Dementia (PrAISED). PLOS ONE 2019; 14(5): e0217387.
429 https://doi.org/10.1371/journal.pone.0217387

430 33. Palazzo C, Klinger E, Dorner V, et al. Barriers to home-based exercise program
431 adherence with chronic low back pain: Patient expectations regarding new technologies.
432 Ann Phys Rehabil Med 2016; 59(2): 107-113. https://10.1016/j.rehab.2016.01.009

433 34. Lansing RW, Gracely RH, Banzett RB. The multiple dimensions of dyspnea: review
434 and hypotheses. *Respir Physiol Neurobiol* 2009; 167(1):53–60.
435 https://doi.org/10.1016/j.resp.2008.07.012

436 35. Mills HL, Higgins JPT, Morris RW, et al. Detecting Heterogeneity of Intervention
437 Effects Using Analysis and Meta-analysis of Differences in Variance Between Trial
438 Arms. *Epidemiol* 2021; 32(6):846–54. https://doi.org/10.1097/EDE.0000000000001401

439 36. Sudre CH, Murray B, Varsavsky T, et al. Attributes and predictors of long COVID. Nat
440 Med 2021: 27(4): 626-631

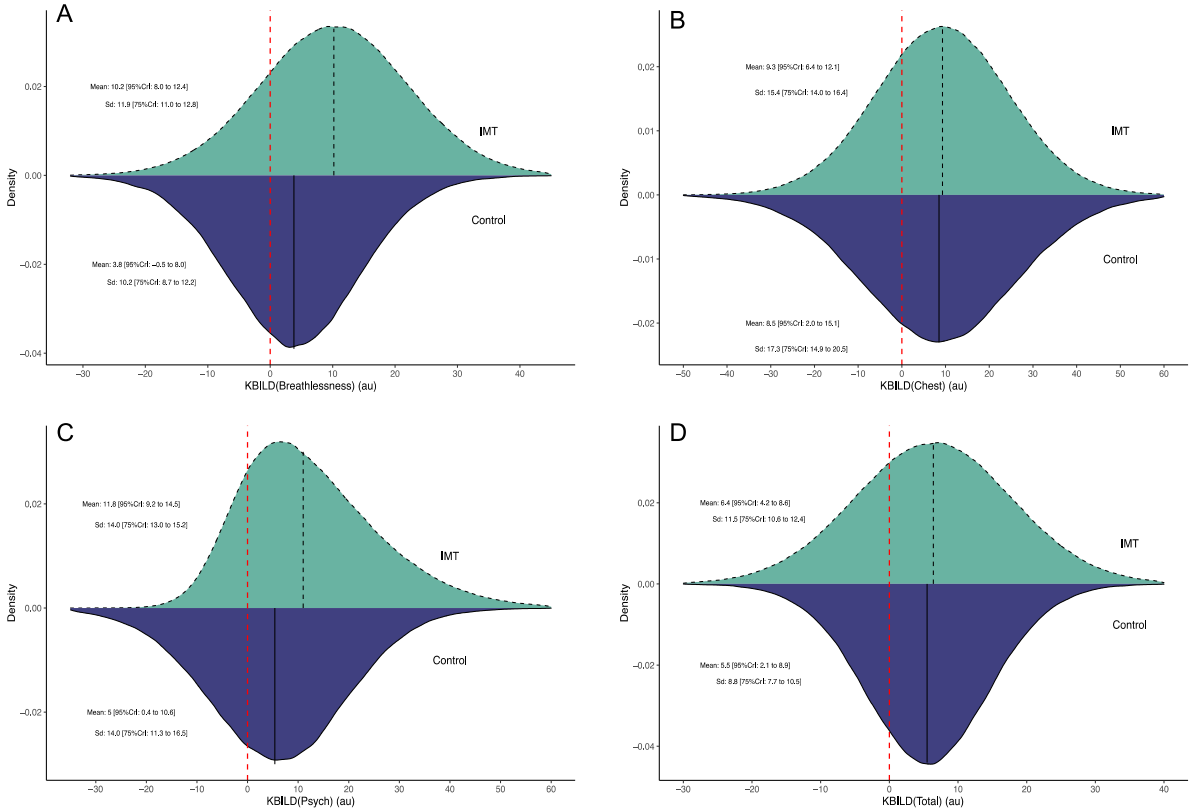


**Figure 1:** Distribution of change scores in KBILD Breathlessness (A), Chest (B), Psychological (C) and total (D) scores following IMT (green) and control (blue). Black vertical lines represent the estimated mean changes, and the dashed red line represents zero. KBILD: 15-item Kings Brief Interstitial Lung Disease; CrI: Credible interval.

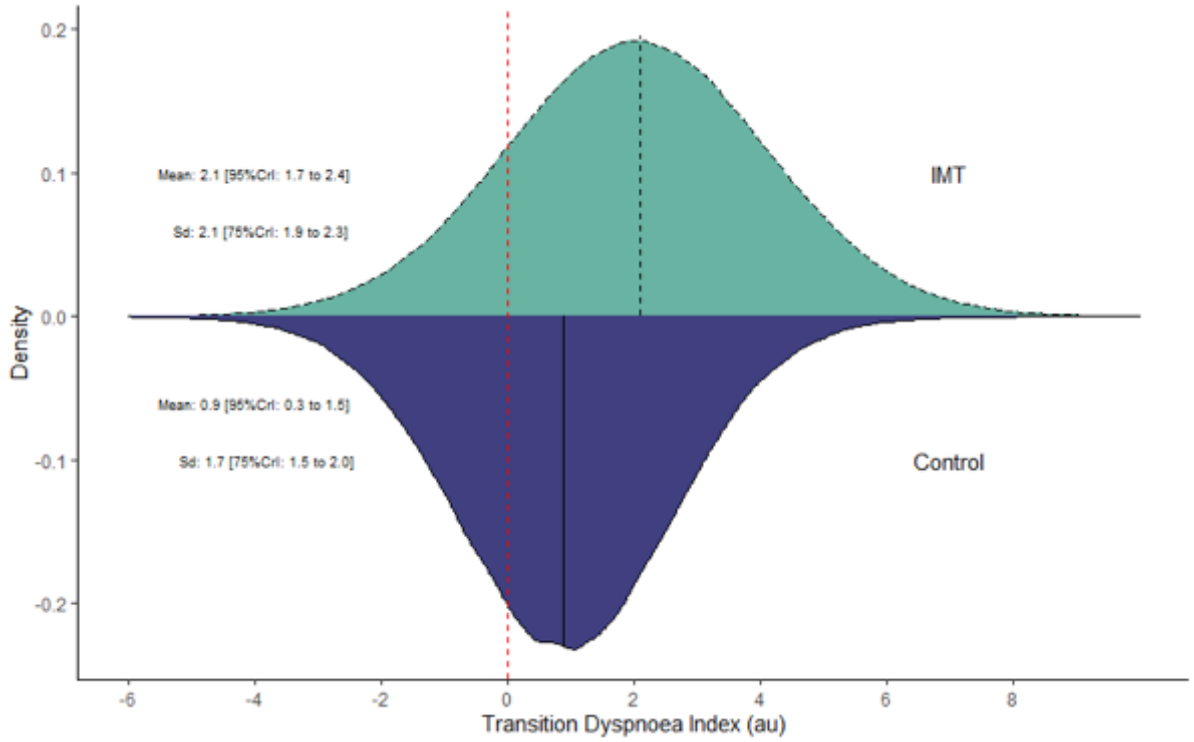


**Figure 2:** Distribution of Transition Dyspnoea Index scores following IMT (green) and control (blue). Black vertical lines represent the estimated mean changes, and the dashed red line represents zero. CrI: Credible interval.

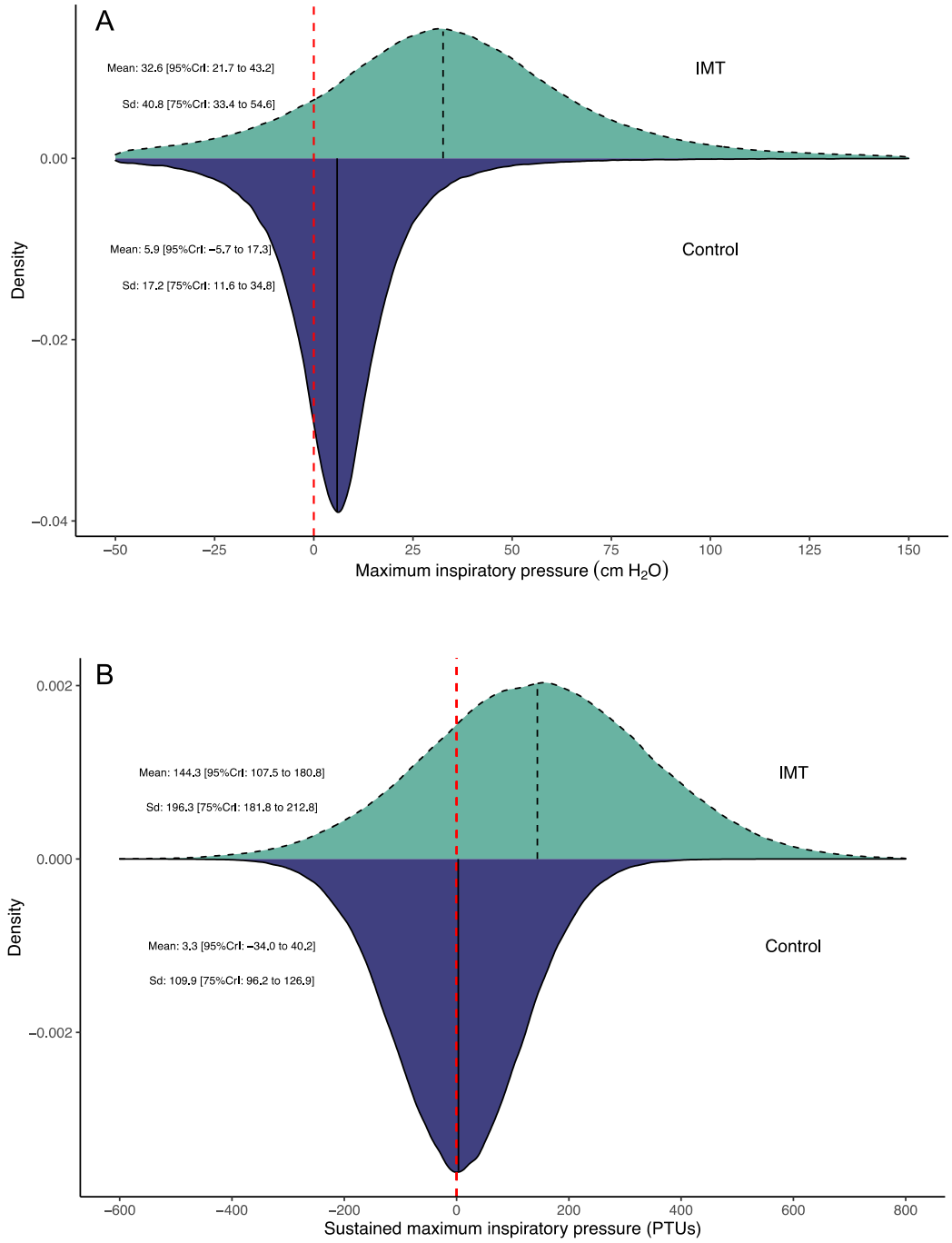


**Figure 3:** Distribution of change scores for maximal inspiratory pressure (A) and sustained maximal inspiratory pressure (B) following IMT (green) and control (blue). Black vertical lines represent the estimated mean changes, and the dashed red line represents zero. CrI: Credible interval.

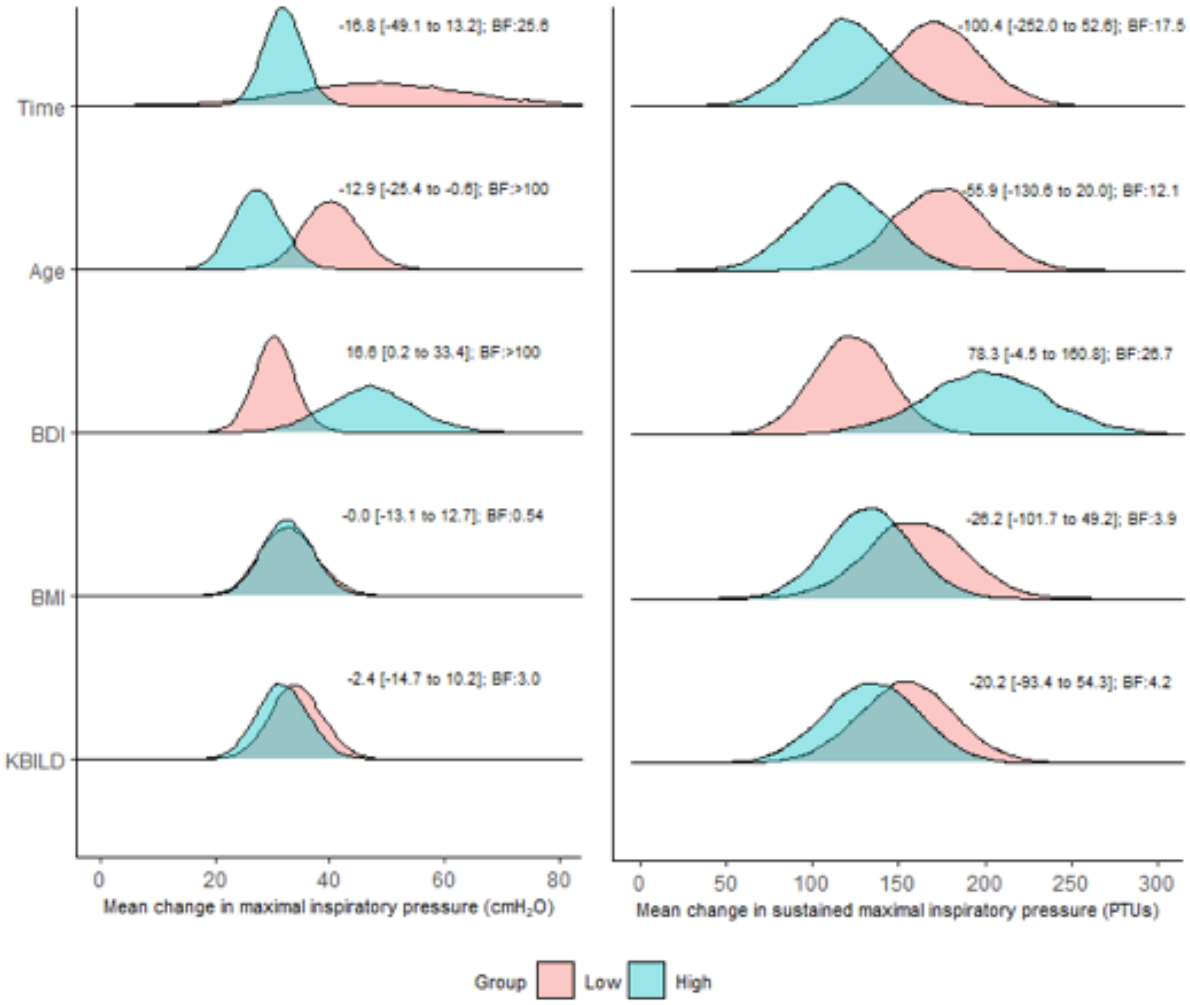


**Figure 4:** Density plots illustrating subgroup analyses of dichotomised participant characteristics exploring relative treatment effect modification for changes in maximal inspiratory pressure (left) and sustained maximal inspiratory pressure (right). Values and credible intervals provided estimate the difference in mean change following training between participants in the high relative to low group (positive values denote greater mean change in the high group). Time: time since COVID [low: ≤ 3months; high: >3 months]; Age [low:<50 years; high: ≥50 years]; BMI: Body Mass Index [low: <25 kg•m$^{-2}$; high: ≥25 kg•m$^{-2}$]; age [low:<50 years; high: ≥50 years] KBILD: 15-item Kings Brief Interstitial Lung Disease baseline total score [low: <53; high: ≥53]; BDI: Baseline Dyspnoea Index [low: ≤6 units; high: >6 units]); BF: Bayes factor.

**Table 1** Participant characteristics

| | IMT (*n*=111) | Control (*n*=36) |
|---|---|---|
| Males / Females | 19 / 92 | 2 / 34 |
| Age (y) | 48 (11) | 49 (12) |
| BMI ($kg•m^{-2}$) | 27.8 (6.9) | 27.5 (6.2) |
| Time since COVID-19 (months) | 9.3 (3.6) | 9.4 (3.2) |
| Baseline Dyspnoea Index | 5.8 (2.5) | 5.4 (2.9) |

*Data are shown as mean (SD) unless indicated otherwise.*

**Table 2:** Variation in IMT Intervention Characteristics

| Training Characteristic | Prescribed | Recorded | | |
|---|---|---|---|---|
| | | Median (IQR) | Min | Max |
| Total Sessions (*n*) | 24 | 20 (6) | 0 | 44 |
| Frequency (mean sessions•week$^{-1}$) | 3 | 2.5 (0.750) | 0 | 5.3 |
| Total Breaths (*n*) | 864 | 607 (357.6) | 0 | 1565 |
| Mean Breaths•Session$^{-1}$ (*n*) | 36 | 33.2 (8.35) | 0 | 37.9 |
| Mean Breath Duration (secs) | - | 11.4 (4.6) | 0 | 22.2 |
| Cumulative Power (PTUs) | - | 222,302 (187,338) | 0 | 602,809 |
| Cumulative Power: Baseline MIP | - | 3,152 (2,867) | 0 | 16,047 |
| Cumulative Power: Baseline SMIP | - | 525 (483) | 0 | 2,245 |

SD: standard deviation; PTUs: pressure time units; MIP: maximal inspiratory pressure; SMIP: sustained maximal inspiratory pressure. IQR: Interquartile Range. Cumulative Power: Baseline MIP and Cumulative Power: Baseline SMIP were calculated by dividing Cumulative Power (PTUs) by baseline MIP and SMIP respectively.

**Table 3:** Assessment of heterogeneous treatment effects across dependent variables based on group change scores

| Variable | Mean difference [95% CrI] | Standard deviation difference [75% CrI] | Bayes factor | Distribution | Proportion of response [95% CrI] |
|---|---|---|---|---|---|
| KBILD (Breathlessness) (au) | 6.4 [-0.3 to 13.2] | 1.7 [-1.3 to 4.3] | 0.65 | Normal distribution | NA |
| KBILD (Psychological) (au) | 6.8 [-1.2 to 15.1] | -0.0 [-3.9 to 3.3] | 0.41 | Normal distribution | NA |
| KBILD (Chest) (au) | 0.8 [-9.6 to 10.9] | -1.9 [-6.6 to 1.9] | 0.54 | Normal distribution | NA |
| KBILD (Total) (au) | 0.9 [-4.9 to 6.6] | 2.7 [0.2 to 4.9] | 1.49 | Normal distribution | NA |
| Transition Dyspnoea Index (au) | 1.1 [0.2 to 2.1] | 0.4 [-0.1 to 0.8] | 0.94 | Normal distribution | NA |
| Maximum inspiratory pressure (cm $H_2O$) | 26.6 [10.5 to 42.7] | 22.8 [4.7 to 37.7] | >100 | T-distribution | 0.84 [0.63 to 1.0] |
| Sustained maximum inspiratory pressure (PTUs) | 141.2 [68.0 to 42.7] | 86.8 [55.7 to 116.7] | >100 | Normal distribution | 0.95 [0.76 to 1.0] |

Mean difference: Difference in the mean change score between IMT and control groups. Standard deviation difference: Difference in the standard deviation of change scores between IMT and control groups. Bayes factors: Values greater than 1 provide evidence for difference in the standard deviation of change scores between IMT and control groups. KBILD: 15-item King's Brief Interstitial Lung Disease. CrI: Credible interval. NA: Not applicable due to no clear evidence of heterogenous treatment effects.

**Table 4:** Assessment of relative treatment effect modification of training-related variables for changes in maximal and sustained maximal inspiratory pressure

| Variable | Maximal inspiratory pressure (MIP) | | Sustained maximal inspiratory pressure (SMIP) | |
|---|---|---|---|---|
| | $\beta$ Change Score [95%CrI:] | Bayes Factor | $\beta$ Change Score [95%CrI:] | Bayes Factor |
| Number of sessions (*n*) | 2.7 [-3.0 to 8.7] | 6.5 | 5.6 [-28.8 to 39.0] | 5.9 |
| Total breaths (*n*) | 5.3 [-0.4 to 11.3] | 37.8 | 15.7 [-17.1 to 50.1] | 43.2 |
| Mean breaths per session (*n*) | 8.3 [2.5 to 14.0] | >100 | 26.1 [-7.9 to 59.4] | 96.9 |
| Breath duration (s) | 4.8 [-1.7 to 11.1] | 25.4 | 25.8 [-7.4 to 60.1] | 85.1 |
| Cumulative Power (PTUs) | 10.9 [5.3 to 16.8] | >100 | 63.7 [32.2 to 95.3] | >100 |
| Cumulative Power: Baseline MIP | 14.3 [7.9 to 21.0] | >100 | 57.9 [25.0 to 90.6] | >100 |
| Cumulative Power: Baseline SMIP | 9.8 [4.4 to 15.5] | >100 | 90.2 [61.4 to 120.0] | >100 |

Training variables were standardised such that $\beta$ represents the expected increase/decrease in the dependent variable change scores for a standard deviation increase in the training variable. CrI: Credible interval; PTUs: pressure time units. Cumulative Power: Baseline MIP and Cumulative Power: Baseline SMIP were calculated by dividing Cumulative Power (PTUs) by baseline MIP and SMIP respectively.