SportRxiv

PREPRINT - NOT PEER REVIEWED

The effects of lengthened-partial range of motion resistance training of the limbs on arm and thigh muscle cross-sectional area: a multi-site cluster trial

Luke Carlson[1], David Gschneidner[1], James Steele[2,3], James Fisher[3*]

[1]Discover Strength
[2]Steele Research Limited
[3]Department of Sport and Health, Solent University

*Correspondence: james.fisher.phd@outlook.com

Cite as: Luke Carlson et al. (2024). The effects of lengthened-partial range of motion resistance training of the limbs on arm and thigh muscle cross-sectional area: a multi-site cluster trial. *SportRxiv.* DOI: https://doi.org/10.51224/SRXIV.485
Supplementary Materials: https://osf.io/t9auy/

Abstract

It has recently been hypothesised that resistance training (RT) performed using a partial range of motion (ROM) at long muscle lengths (i.e., "lengthened partials") might optimise gains in muscle size. Improvements in muscle size however are typically small, even smaller in trained people due to the linear-logarithmic adaptation to RT over time, and thus between intervention differences in effects are likely to be very small. As such, in contrast to most studies in the field which aim to detect differences between interventions, we sought to conduct a highly powered pre-registered test of the statistical equivalence of two RT interventions in previously trained participants; namely full ROM (fROM) and "lengthened partial" ROM (lpROM). A randomised controlled cluster trial across 15 sites was employed. Our primary outcome was hypertrophy operationalised as muscle cross sectional area (CSA) estimated from circumference and skinfold measurements of the upper arm and thigh. Secondary outcomes were strength operationalised as estimated one-repetition maximum estimated from submaximal repetitions to momentary failure for chest press, leg press, and pulldown machines. At the participant level, participants were randomly assigned to either the lpROM (n = 163) or fROM (n = 134) RT intervention condition. Participants underwent pre-intervention testing and then participated in a 12-week intervention with post-intervention testing following this. Our primary estimand of interest was the condition by time interaction effect from our pre-registered analysis reflecting the standardised between condition difference in change in hypertrophy over time. The estimate for this effect for the arm estimated muscle CSA was -0.032 [95%CI: -0.123, 0.058] and for the thigh estimated muscle CSA was 0 [95%CI: -0.094, 0.094]. The p-values for equivalence were p=0.071 for the arm muscle, and p=0.019 for the thigh muscle. As such, considering our inference criteria with alpha set at 0.01 and adjusted to 0.005 for multiple outcomes, we were unable to reject the null hypothesis that the condition:time interaction effect was outside of the SESOI [-0.1, 0.1]. Main effects for time were small for estimated muscle CSA and also strength, with both being in line with prior predictions from theoretical linear-log growth models. Additional exploratory analysis suggest that both the main effects of time, and any interaction effects for condition by time, are likely small. These findings are in line with other recent evidence regarding the comparison of fROM and lpROM specifically and suggest that between condition effects are likely small and practically equivalent. More broadly, this study highlights that the effects of RT in trained persons should be expected to be small and that current studies in the field of RT are woefully underpowered to be able to detect their effects, let alone test between intervention comparisons.

Keywords: *resistance training; range of motion; hypertrophy; strength; trained participants*

All authors have read and approved this version of the manuscript.
The manuscript was last updated on December 13, 2024

Introduction

Resistance training (RT) is the most widely supported intervention for increasing muscle size (i.e., hypertrophy) in adults. The effects of manipulating RT protocol variables to optimise hypertrophy continues to be a key topic of interest in the field (B. Schoenfeld et al., 2021). One such variable is the range of motion (ROM) used, and this has been a current source of debate in the field with several recent systematic reviews and meta-analyses examining its effects (Kassiano et al., 2023; Pallarés et al., 2021; B. J. Schoenfeld & Grgic, 2020; Wolf et al., 2023).

A specific hypothesis that has gained attention from this area of debate is that RT performed with a partial ROM at a long muscle lengths (so called "lengthened-partials") might produce greater hypertrophy than when RT is performed using a full ROM (Kassiano et al., 2023; Wolf et al., 2023). Indeed, in the recent meta-analysis from Wolf et al. (2023) whilst partial and full ROM RT were found to be on average similarly effective for increasing muscle size (standardised mean effect 0.04 [95% quantile interval: -0.17, 0.25]), when distinguishing between studies performing partial ROM at either shorter or longer muscle lengths those using the latter tended towards greater hypertrophy in partial ROM (standardised mean effect of -0.28 [95% quantile interval: -0.81, 0.16]), though this was quite imprecise as a result of the number of studies included (k = 8). In a recent critical narrative review Moreno et al. (2024) argued that conclusions regarding the superiority of partial ROM training at long muscle lengths may be premature given the state of the data, that there may be region specific effects with greater hypertrophy at more distal muscle sites when using partial ROM at long muscle lengths, and that any effects require replication across different muscle groups before claiming generalisability. However, a more recent meta-analysis by Varovic et al. (2024) (k = 12) specifically examined the effects of mean muscle length during RT upon regional measurements of muscle size. Building upon the previous work of Wolf et al. (2023) utilising Bayesian models they found trivial standardised mean effects of longer compared to shorter mean muscle lengths during RT at 25% (0.04 [95%QI: -0.07, 0.15]), 50% (0.07 [95%QI: -0.02, 0.15]), and 75% (0.09 [95%QI: -0.0, 0.19]) muscle sites.

One issue with the current debate regarding the comparative effects of partial ROM at long muscle lengths compared to full ROM RT[1] is that, as is typical for most RT studies (Steele, Fisher, Smith, et al., 2023), sample sizes are small and thus most studies are underpowered and imprecise statistically to be able to make strong claims one way or another regarding the hypothesis being tested. Indeed, the average intervention effect over time for a typical ~12-week RT intervention compared with non-training controls in untrained participants is a standardised mean effect of 0.34 [95%CI: 0.29, 0.39] (Steele, Fisher, Smith, et al., 2023). Effects this size for a condition by time interaction (i.e., the difference in difference between RT and non-training controls over time[2]) imply that any further interaction effect (i.e., the comparative difference in difference between one RT intervention and another RT intervention

[1] Or indeed any debates regarding the comparative effectiveness of different RT interventions manipulating training variables.

[2] Though an interaction effect estimated from the comparison of RT and non-training controls over time, given the comparison conditions this reflects the *main* causal effect of time for an RT intervention.

over time) should necessarily be smaller assuming that both are at least effective over time. Confounding this further is the now well supported fact that adaptation to RT over time follows a roughly linear-log growth function (Latella et al., 2024; Steele, Fisher, Giessing, et al., 2023) (see also https://osf.io/7fdvq and https://osf.io/xrfyb). Given that the main effect of RT over time is diminished in trained participants, the interaction effect for any between RT condition comparisons is likely to be even smaller than in untrained participants. In fact, a recent study (Wolf et al., 2024) in trained participants found standardised mean effects for condition by time interaction (calculated from their raw open data: https://osf.io/q9djw) ranging from -0.01 [95%CI: -0.12, 0.09] to 0.1 [95%CI: -0.02, 0.22]. In this study a within-participant design was utilised which improves statistical power and precision to detect effects (MacInnis et al., 2017), yet their estimates reinforce how small the differences in effectiveness of any given RT interventions are likely to be.

Given these factors, whilst there may really be a real difference in effectiveness between interventions, it seems unlikely that any between intervention differences would reach any reasonable smallest effect size of interest (SESOI). Thus, studies intending to make strong claims regarding the presence of such an effect would need to be prohibitively large in sample, and likely so even if utilising other methods for improving power and precision such as within participant designs and high frequency measurement (MacInnis et al., 2017; Swinton, 2024). Instead, it makes sense to examine hypotheses relating to the equivalence of interventions as has recently been argued for in the sport and exercise sciences (Mazzolari et al., 2022)). As such, in contrast to most studies in the field, we sought to conduct a highly powered pre-registered test of the equivalence of two RT interventions in previously trained participants; namely full ROM and "lengthened partial" ROM.

Methods

Experimental approach to the problem

A randomised controlled cluster trial across multiple sites was employed. At the participant level, participants were randomly assigned to either the lengthened partial ROM (lpROM) or full ROM (fROM) RT intervention condition. Participants underwent two baseline testing time points (t_{-1} and t_0) and then participated in a 12-week intervention with post testing following the 12-week intervention (t_1). This study was pre-registered July 2024, available at https://osf.io/9sgjk.

Participants and Sample Size

Our pre-registered target sample size was 300 participants (across 15 Discover Strength sites, target of 20 participants recruited at each site). The full details of the sample size justification for statistical power including simulations, assumptions, analyses, and inference criteria can be seen in the pre-registration: https://osf.io/9sgjk. Briefly, we assumed the aforementioned theoretical linear-log growth function and estimated the expected main effect of time (standardised mean difference [SMD] of ~0.05) for participants with the prior training experience

required for inclusion noted below and assumed any condition by time interaction effect to be at most half of this (i.e., ~0.025). We considered this to be within what we felt to be the SESOI for changes in muscle size; an SMD ranging [-0.1, 0.1] which we also corroborated by consulting other researchers in the area[3]. We then simulated to determine statistical power at different sample sizes assuming these effects and testing for equivalence following the analysis plan detailed below.

Following ethical approval for this study by Solent University Health, Exercise, and Sports Science Research Ethics and Innovation Committee (reference number: fishj1HESS2024), 298 participants were recruited from 15 locations of Discover Strength personal training studio (USA). All participants were existing clients with a minimum of 6 months RT experience with Discover Strength (though may have also had prior training experience before joining Discover Strength as members) ensuring all participants were familiar with supervised, high effort (i.e., training to momentary failure, and occasionally the use of the use of advanced training techniques such as drop-sets, pre- or post-exhaustion, forced repetitions, etc.), low-volume (i.e., a single set of each exercise), and twice-weekly training practices. Participants were instructed not to (and confirmed that they did not), engage in any muscle strengthening exercise outside of their supervised strength training sessions. Participants were also asked to maintain normal dietary patterns and daily activities or other physical activities, exercise, and sports that they currently participated in (e.g., not to begin additional exercise strategies to enhance weight loss/gain). All participants signed an informed consent form prior to any data collection.

Muscle Area Measurement

The primary outcomes were arm- and thigh- muscle area estimated from anthropometric measurements. Estimates of both arm muscle cross sectional area (CSA) and thigh muscle CSA were made from anthropometric measurements using methods and equations described previously by Heymsfield, et al. (1982) and Housh et al. (1995), respectively. For arm muscle CSA, circumference was measured at the midpoint between the tip of the acromion and the olecranon process with the arm hanging relaxed by the participant's side and taking the triceps skinfold as a vertical fold at the same point using callipers. For thigh muscle CSA, circumference was measured at the midpoint of the inguinal crease and the proximal border of the patella, and a thigh skinfold taken as a vertical fold at the same point using the same callipers. Measurements were taken by multiple instructors across multiple locations. All staff at each site conducting anthropometric measurements underwent initial training collecting measurements from at least 5 different people over two occasions, and in addition two baseline measurement occasions were used to provide further practice. For each muscle CSA outcome three measurements were taken for both circumference and skinfold at each time point and all were used in analysis. This choice was based on our simulations (see https://osf.io/9sgjk) regarding the impact of multiple measurements at each time point, whilst taking into account estimated measurement error determined from prior data, upon statistical power.

[3]Note, we should also add that the same smallest effect size of interest was elicited by the authors of a recent meta-analysis, including JS, examining the effects of muscle length upon regional hypertrophy (Varovic et al., 2024). Indeed, prior to its conduct we had consulted some of the authors of the meta-analysis to elicit the SESOI.

Strength measurement

For our strength measurements we deviate from our pre-registration. However, given this was explicitly pre-registered as an exploratory outcome and the deviations increase the precision of our estimates, we felt this deviation was justifiable. We originally intended to measure strength as a secondary outcome specifically at t_0 and t_1 i.e., pre- and post-intervention, using an estimated 10 repetition maximum (RM) test from which 1RM would be estimated for the leg-press, chest-press, and pulldown exercises performed as indicated in the pre-registration. However, we realised that the system used by Discover Strength to record client workouts, StrengthPortal https://strengthportal.com/, allowed us to instead track every single set of each exercise performed and, combined with the single set protocol our intervention utilised, we were instead able to extract and model estimated 1RM from the loads and repetitions performed for all leg-press, chest-press, and pulldown machine exercises by each participant over the entire duration of the intervention periods. This meant that instead of merely pre- and post-intervention strength outcomes we had a far greater number of strength outcomes which could be modelled across the intervention period. Approaches such as this to utilise high frequency outcome measurement have recently been recommended for RT research to increase statistical power considerably even in the face of possible measurement error increases with estimation methods such as submaximal load RM tests (Swinton, 2024). Thus, the loads lifted, and number of repetitions performed, were used to estimate 1RM using the Baechle (2008) equation: predicted 1RM = load lifted x (1 + [0.033 × number of repetitions]). We considered that this method provides strong ecological validity to realistic training conditions, indeed this approach actually utilised realistic training conditions, because most people infrequently test or use their maximal strength. MF during testing as such was defined similarly to the RT intervention as the point at which, despite the greatest effort, the participant failed to complete the concentric phase of a repetition (Steele et al., 2017).

Intervention

Training was performed 2×/week (with at least 48 hours between sessions) using two different workouts (see Table 1). These exercises were chosen to present two whole-body workouts covering all major muscle groups in each workout with particular attention to the upper arms and thighs. All exercises were performed for a single set (unless stated otherwise in Table 1) and all participants used the same approximate relative load. Thus, the two training intervention conditions were the same in all variables with the exception of the ROM used for the single-joint exercises for the arms and legs: the lpROM condition performed only the first half of the range of motion where the muscle begins in its most lengthened position for that exercise. The fROM condition performed full ROM training for the same volume and relative load (e.g., 6-8RM) as the lengthened partial ROM condition. Furthermore, each muscle action was performed for the same repetition duration, i.e., a 2second concentric: 4second eccentric regardless of ROM. This meant the lpROM participants effectively moved at half the speed of the full ROM group since they covered half the distance in the same time. All workouts were supervised, and all exercises were completed to MF as defined above.

Table 1: Workouts performed for the intervention.

Workouts
Session 1
1. Seated calf press
2. Leg press - mid seat position
3. Adductor
4. Leg extension
5. Leg curl
6. Leg extension
7. Leg curl
8. Chest press
9. Pulldown - pronated grip
10. Overhead press
11. Pullover
12. Incline supinated dumbbell curl
13. French press tricep extension
14. Incline hammer curl
15. Core torso rotation
Session 2
1. Leg press - upright seat position
2. Abductor
3. Leg extension
4. Leg curl
5. Leg extension
6. Leg curl
7. Tibia
8. Lateral raise
9. Incline press
10. Seated row
11. Overhead press
12. Incline supinated dumbbell curl
13. French press tricep extension
14. Incline hammer curl
15. Abdominals
16. Core lumbar spine

Note:
ROM = Range of motion; Bold exercises performed with either full ROM, or lengthened partial ROM

Statistical Analysis

All code utilized for data preparation and analyses are available in either the Open Science Framework page for this project https://osf.io/t9auy/ or the corresponding GitHub repository

https://github.com/jamessteeleii/lenthened_partial_trial. We cite all software and packages used in the analysis pipeline using the `grateful` package (Rodriguez-Sanchez et al., 2023) which can be seen here: https://osf.io/wrzgf.

Primary Pre-registered Analysis (Hypertrophy - Arm and Thigh Estimated Muscle Cross Sectional Area)

As noted, the project was previously pre-registered including the analysis plan, model to be employed, parameter of primary interest, and specific hypothesis relating to this. The full details of this including the derivation of our hypotheses from prior evidence and theory regarding the expected effects of resistance training upon hypertrophy are fully detailed in the pre-registration for the reader (see https://osf.io/9sgjk). Here we reiterate our primary hypothesis related to the between condition comparison of lengthened partial ROM vs. full ROM resistance training interventions on arm and thigh estimated muscle cross sectional area (CSA) over time i.e., the time-by-condition interaction, where time is pre- and post-intervention (i.e., T0 and T1) and condition is the two aforementioned interventions. We tested for the equivalence of the slopes for time of lengthened partial ROM against the full ROM comparator with a smallest effect size of interest of 0.1 SMD (see pre-registration for full justification for our choice of smallest effect size of interest). Our hypothesis is thus that the lengthened partial ROM resistance training condition will produce changes over time in our primary outcome measure of estimated muscle size (measured for both arm and thigh with alpha corrections for multiple outcomes; see below) that are not larger (or smaller) than the smallest effect size of interest when compared with the full ROM resistance training intervention condition. More specifically:

H0: The interaction effect for time (pre- and post-intervention) and condition (lengthened partial ROM or full ROM resistance training intervention) for muscle CSA will differ from the smallest effect size of interest - upper and lower bound of confidence interval for between condition effect will be outside of or include the upper or lower limits of smallest effect size of interest i.e., [-0.1,0.1].

H1:The interaction effect for time (pre- and post-intervention) and condition (lengthened partial ROM or full ROM resistance training intervention) for muscle CSA will be equivalent to the smallest effect size of interest - upper and lower bound of confidence interval for between condition effect will be inside the upper or lower limits of smallest effect size of interest i.e., [-0.1,0.1].

A linear mixed effects model was fit using the `lme4` package and using Restricted Maximum Likelihood estimation for both arm and thigh estimated muscle CSA outcomes with fixed effects for time, condition, and condition:time interaction (our estimate of interest), random intercepts for both site id and participant id, and a random slope for time within participant clusters. The model equation was as follows:

$$
\begin{aligned}
\mathsf{y}_i &\sim N\left(\alpha_{j[i],k[i]} + \beta_{1j[i]}(\text{time}), \sigma^2\right) \\
\begin{pmatrix} \alpha_j \\ \beta_{1j} \end{pmatrix} &\sim N\left(\begin{pmatrix} \gamma_0^{\alpha} + \gamma_1^{\alpha}(\text{condition}) \\ \gamma_0^{\beta_1} + \gamma_1^{\beta_1}(\text{condition}) \end{pmatrix}, \begin{pmatrix} \sigma^2_{\alpha_j} & \rho_{\alpha_j\beta_{1j}} \\ \rho_{\beta_{1j}\alpha_j} & \sigma^2_{\beta_{1j}} \end{pmatrix}\right), \text{ for participant j = 1, ... ,J} \\
\alpha_k &\sim N\left(\mu_{\alpha_k}, \sigma^2_{\alpha_k}\right), \text{ for site k = 1, ... ,K}
\end{aligned}
$$

Where y_i was the estimated muscle CSA for either arm or thigh standardised following the approach of Penney (2023) i.e., using the error term of a simple linear model including only the randomised condition predictor avoiding possible attenuation issues with using z-scores. Condition was coded as centred (i.e., fROM = -0.5; lpROM = 0.5) such that the main effect of time was interpretable as the mean effect across both conditions.

Our primary test for equivalence was upon the condition:time effect i.e., $\gamma_1^{\beta_1}(\text{condition})$ from our primary pre-registered model using the `marginaleffects` packages `hypotheses()` function against our SESOI of 0.1. We pre-registered an alpha of 0.01, corrected to 0.005 given we had two primary outcomes relating to the same construct (i.e., muscle hypertrophy) to draw inferences regarding equivalence. We also examined the estimated main time effect i.e., $\beta_{1j[i]}(\text{time})$ in our model descriptively in terms of it's magnitude and precision, and for visualisation present the un-pooled linear predictions for each participant on the raw scale (i.e., cm^2) in addition to the means and 95% quantile intervals for the raw pre- and post-intervention data.

Secondary Exploratory Analyses

Strength

We originally noted in our pre-registration that strength would be measured both pre- and post-intervention using an estimated 10RM from which 1RM would be predicted for the leg-press, chest-press, and pulldown machine exercises. However, as noted above, because of the utilisation of the strength portal to manage and record all workouts by Discover Strength combined with the single set protocol we were instead able to extract and model estimated 1RM from the loads and repetitions performed for all leg-press, chest-press, and pulldown machine exercises by each participant over the entire duration of the intervention periods. Approaches such as this to utilise high frequency outcome measurement have recently been recommended for resistance training research to increase statistical power considerably even in the face of possible measurement error increases with estimation methods such as submaximal load RM tests (Swinton, 2024).

A linear mixed effects model was fit using the `lme4` package and using Restricted Maximum Likelihood estimation for leg-press, chest-press, and pulldown estimated 1RM outcomes with fixed effects for time, condition, and condition:time interaction, random intercepts for site id, participant id, and exercise nested within participant id[4] and a random slope for time within

[4]Note, we grouped outcomes by machine type i.e., leg-press, chest-press, and pulldown, but because the specific machines differed in some cases between sites, and in some cases participants trained utilising different exercises

participant and exercise nested within participant clusters. The model equation was as follows:

$$
\begin{aligned}
\mathrm{y}_i &\sim N\left(\mu, \sigma^2\right) \\
\mu &= \alpha_{j[i],k[i],l[i]} + \beta_{1j[i],k[i]}(\text{time}) + \beta_2(\text{condition}) + \beta_3(\text{condition} : \text{time}) \\
\begin{pmatrix} \alpha_j \\ \beta_{1j} \end{pmatrix} &\sim N\left(\begin{pmatrix} \mu_{\alpha_j} \\ \mu_{\beta_{1j}} \end{pmatrix}, \begin{pmatrix} \sigma^2_{\alpha_j} & \rho_{\alpha_j\beta_{1j}} \\ \rho_{\beta_{1j}\alpha_j} & \sigma^2_{\beta_{1j}} \end{pmatrix}\right) \text{, for exercise:participant j = 1, ... ,J} \\
\begin{pmatrix} \alpha_k \\ \beta_{1k} \end{pmatrix} &\sim N\left(\begin{pmatrix} \mu_{\alpha_k} \\ \mu_{\beta_{1k}} \end{pmatrix}, \begin{pmatrix} \sigma^2_{\alpha_k} & \rho_{\alpha_k\beta_{1k}} \\ \rho_{\beta_{1k}\alpha_k} & \sigma^2_{\beta_{1k}} \end{pmatrix}\right) \text{, for participant k = 1, ... ,K} \\
\alpha_l &\sim N\left(\mu_{\alpha_l}, \sigma^2_{\alpha_l}\right) \text{, for site l = 1, ... ,L}
\end{aligned}
$$

Where y_i was the estimated 1RM on each exercise for either chest press, leg press, or pulldown machines standardised following the approach of Penney (2023). Condition was coded as centred (i.e., fROM = -0.5; lpROM = 0.5) such that the main effect of time was interpretable as the mean effect across both conditions. In addition, time was originally in days across the intervention but in the model was scaled to a period of 84 days (i.e., 12 weeks) such that the main effect for time was interpretable as the effect over that period of time.

We examined both the estimated main time effect i.e., $\beta_{1j[i],k[i]}(\text{time})$, and condition:time interaction effect i.e., $\beta_3(\text{condition:time})$, in this model descriptively in terms of their magnitude and precision, and for visualisation also present un-pooled predictions using penalised cubic spline smooths for each participant on the raw scale (i.e., kilograms) in addition to the group level predictions using penalised cubic spline smooths for each condition using raw strength data over time.

Further, we refit the linear-log robust multilevel meta-regression fit to data from Steele et al. (2023) for hypertrophy outcomes during the planning and pre-registration of this study (see https://osf.io/7fdvq and pre-registration https://osf.io/9sgjk) to the strength outcome data in order to derive the predicted simple training effect over time for participants with at least 6 months (i.e., 24 weeks) prior training experience completing an additional 12 weeks of training i.e., the linear slope for the difference in effect size predicted when moving from 24 to 36 weeks of training time. This was so that we could compare the main effects for time against what might have been *a priori* predicted for this population where we assume a linear-log function to generate strength gains over time with exposure to resistance training.

Hypertrophy

Lastly, our pre-registered model was deliberatively conservative in assuming random slopes for time. Further, though we took two sets of baseline measures (i.e., t_{-1} and t_0), our pre-registration was not wholly clear in whether or not these two measurement timepoints pre-intervention were to be used. As such, our main model above has included only the t_0 mea-

on the machine types (e.g., leg-press performed with seat position higher or lower), we explicitly nested the specific exercise machine ("exercise") within participant id.

surements as pre-intervention as this was what was included in the simulations for sample estimation. Further, we mistakenly included in our pre-registration the condition main effect in our model whereas the mixed effects model version of the analysis of covariance (ANCOVA) approach to analysing data from randomised controlled trial data is a more efficient estimator and better reflects the the study design due to randomisation at participant level prior to intervention exposure but after pre-intervention measurement (see Kurz (2022)).

Considering the relative lack of evidence for *true* inter-individual response variation to resistance training (see Steele et al. (2023) and Robinson et al. *forthcoming* (pre-registration https://osf.io/aw5zx)) we also fit an exploratory model without random slopes or a main effect for condition, and utilising all pre-intervention measurements (i.e., both t_{-1} and t_0), with the intention of presenting more precise estimates of potential effects, particularly for time and time:condition interactions.

A linear mixed effects model was fit using the `lme4` package and using Restricted Maximum Likelihood estimation for both arm and thigh estimated muscle CSA outcomes with fixed effects for time and condition:time interaction, and random intercepts for both site id and participant id. The model equation was as follows:

$$
\begin{aligned}
y_i &\sim N\left(\alpha_{j[i],k[i]} + \beta_1(\text{time}) + \beta_2(\text{condition} \times \text{time}), \sigma^2\right) \\
\alpha_j &\sim N\left(\mu_{\alpha_j}, \sigma^2_{\alpha_j}\right), \text{ for participant j = 1, ... ,J} \\
\alpha_k &\sim N\left(\mu_{\alpha_k}, \sigma^2_{\alpha_k}\right), \text{ for site k = 1, ... ,K}
\end{aligned}
$$

Where y_i was the estimated muscle CSA for either arm or thigh standardised following the approach of Penney (2023) i.e., using the error term of a simple linear model including only the randomised condition predictor avoiding possible attenuation issues with using z-scores. Condition was coded as centred (i.e., fROM = -0.5; lpROM = 0.5) such that the main effect of time was interpretable as the mean effect across both conditions.

We examined both the estimated main time effect i.e., $\beta_1(\text{time})$, and condition:time interaction effect i.e., $\beta(\text{condition:time})$, in this model descriptively in terms of their magnitude and precision. In addition, and similarly to the exploratory strength model, we examined the linear-log robust multilevel meta-regression fit to data from Steele et al. (2023) for hypertrophy outcomes during the planning and pre-registration of this study (see https://osf.io/7fdvq and pre-registration https://osf.io/9sgjk) in order to derive the predicted simple training effect over time for participants with at least 6 months (i.e., 24 weeks) prior training experience completing an additional 12 weeks of training i.e., the linear slope for the difference in effect size predicted when moving from 24 to 36 weeks of training time. This was so that we could compare the main effects for time against what might have been *a priori* predicted for this population where we assume a linear-log function to generate hypertrophy gains over time with exposure to resistance training.

Results

For our primary outcomes of estimated muscle CSA we had sample sizes of n = 134 for fROM and n = 163 for lpROM and, taking into account missing data due to either clerical error or participant dropouts, the number of observations we had at pre-intervention for arm and thigh muscle respectively 393 and 391 for fROM and 472 and 481 for lpROM, and at post-intervention for arm and thigh muscle respectively 331 and 335 for fROM and 406 and 404 for lpROM. For our secondary strength outcomes we had data available for sample sizes of n = 131 for fROM and n = 159 for lpROM and the number of observations we had over the course of the intervention for chest press, leg press, and pulldown machines respectively 1498, 2841, and 1496 for fROM and 1795, 3431, and 1779 for fROM.

Primary Pre-registered Results (Hypertrophy - Arm and Thigh Estimated Muscle Cross Sectional Area)

Our primary estimand of interest was the condition:time interaction effect from our pre-registered analysis reflecting the between condition difference in change in hypertrophy over time. The estimate for this effect for the standardised arm estimated muscle CSA was -0.032 [95%CI: -0.123, 0.058] and for the standardised thigh estimated muscle CSA was 0 [95%CI: -0.094, 0.094]. The *p*-values for equivalence were p=0.071 for the arm muscle, and p=0.019 for the thigh muscle. As such, considering our inference criteria with alpha set at 0.01 and adjusted to 0.005 for multiple outcomes, we were unable to reject the null hypothesis that the condition:time interaction effect was outside of the SESOI [-0.1, 0.1]. Thus we cannot clearly make the claim that there is statistical equivalence between the two interventions in terms of their effects. The raw estimated muscle CSA pre- and post-intervention predictions from un-pooled participant level linear regression in addition to the means and 95% quantile intervals can be see in Figure 1, along with the estimates for the main effect of time and condition:time interaction effect on the standardised scale from the pre-registered model.

Secondary Exploratory Results

Strength

The raw estimated 1RMs at participant and exercise level penalised cubic spline smooths, in addition to the group level, can be seen in Figure 2, along with the estimates for the main effect of time and condition:time interaction effect from the exploratory model. Strength increased with time as seen from the descriptive visualisations of the raw data and from the estimates for the main effect of time from the model, though estimates for the condition:time interaction effect did not appear to indicate any clear difference between conditions in strength gains. The estimates for the main effect of time (chest press = 0.16 [95%CI: 0.141, 0.18]; leg press = 0.188 [95%CI: 0.155, 0.221]; chest press = 0.115 [95%CI: 0.097, 0.133]) were also fairly close to the predicted time effect taken from the linear-log robust multilevel meta-regression fit to data from Steele et al. (2023) i.e., 0.117 [95%CI: 0.105, 0.129].

Figure 1: Primary pre-registered hypertrophy outcomes. The top two panels show the raw estimated muscle CSA pre- and post-intervention predictions from un-pooled participant level linear regression in addition to the means and 95% quantile intervals. The bottom two panels show estimates and 95% confidence intervals on the standardised scale from the pre-registered model for the main effect of time and the condition:time interaction effect.

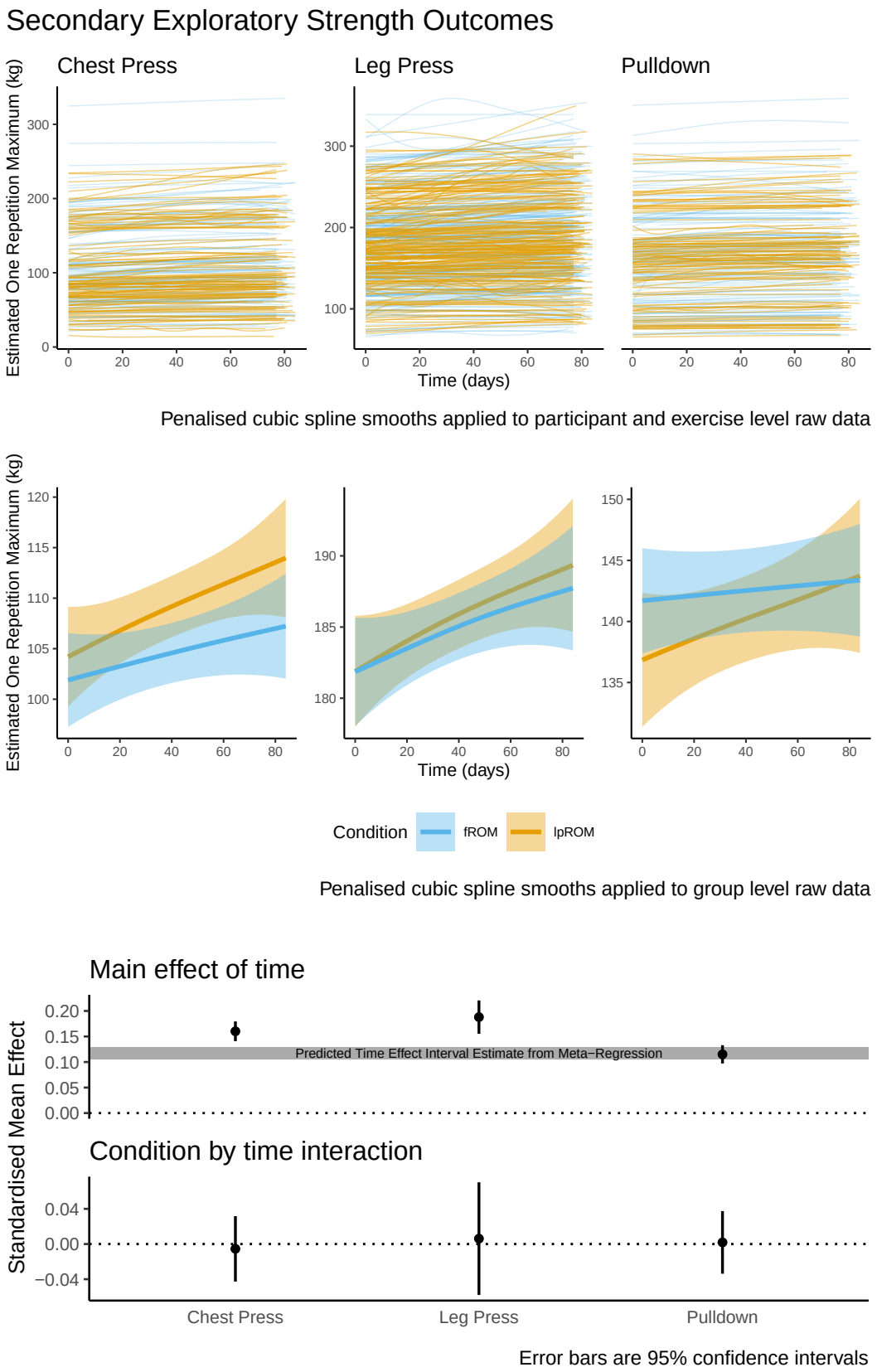

Figure 2: Secondary strength outcomes. The top two rows of panels show the raw estimated 1RMs at participant and exercise level penalised cubic spline smooths, in addition to the group level. The bottom two panels show estimates and 95% confidence intervals on the standardised scale from the exploratory model for the main effect of time and the condition:time interaction effect.

Hypertrophy

The estimates for the main effect of time and condition:time interaction effect from the additional exploratory multilevel ANCOVA model can be seen in Figure 3. As expected the estimates from this model were far more precise than in our main pre-registered model. Though the latter pre-registered did show a main effect estimate of time that excluded zero for both arm muscle 0.08 [95%CI: 0.035, 0.125] and thigh muscle 0.085 [95%CI: 0.038, 0.132] it was not wholly clear or not whether the effects were within the SESOI for hypertrophy. However, from the multilevel ANCOVA model the estimates for time where clearly within the SESOI for both arm muscle 0.034 [95%CI: 0.021, 0.046] and thigh muscle 0.051 [95%CI: 0.037, 0.064] and also fairly close to the predicted time effect taken from the linear-log robust multilevel meta-regression fit to data from Steele et al. (2023) i.e., 0.049 [95%CI: 0.042, 0.057]. Further, with the gain in precision for the estimates in the multilevel ANCOVA model the condition:time effects were more clearly within the SESOI.

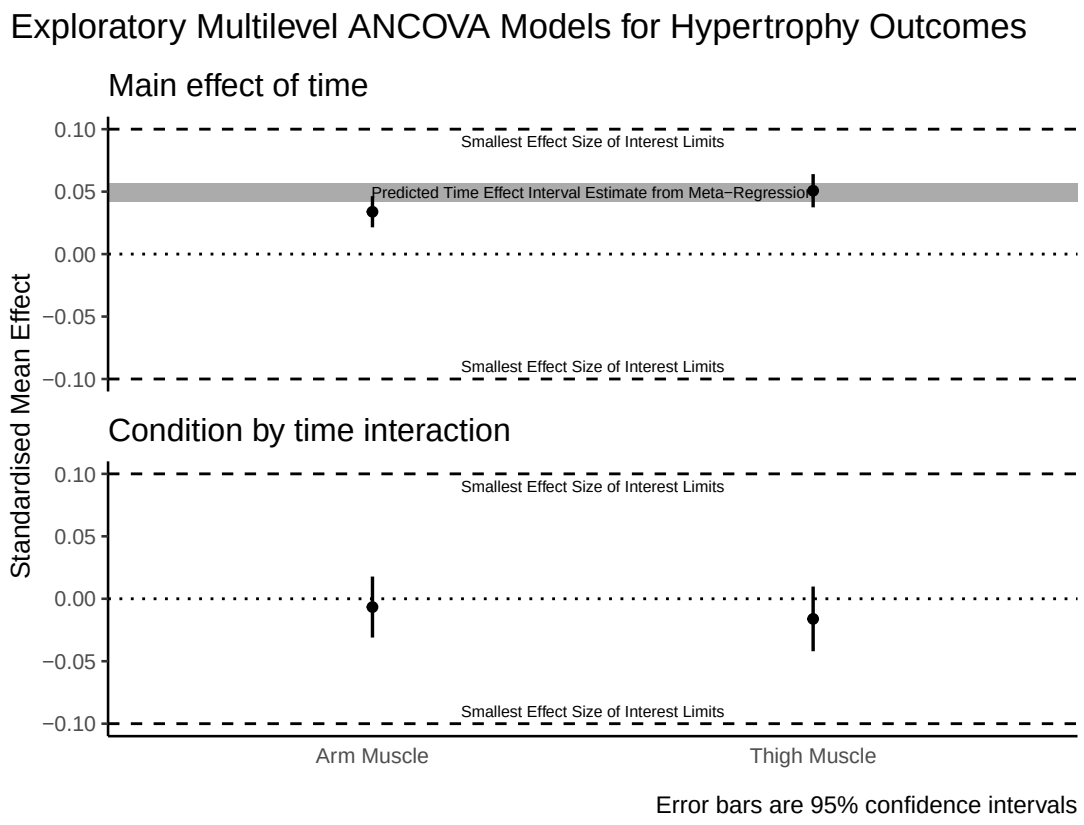

Figure 3: Exploratory multilevel analysis of covariance (ANCOVA) mode for hypertrophy outcomes. The two panels show estimates and 95% confidence intervals on the standardised scale from the pre-registered model for the main effect of time and the condition:time interaction effect.

Discussion

To our knowledge this study is the first to have attempted a severe test (i.e., highly powered, pre-registered) of a theoretically derived prediction whilst also considering the practical meaningfulness (i.e., setting a SESOI) of two RT interventions. In a large-scale multi-site randomized trial we compared the effects of RT using a fROM or lpROM upon muscle size in trained participants and tested for equivalence. The p-values for equivalence were p=0.019 for the arm muscle, and p=0.071 for the thigh muscle. As such, considering our inference criteria with alpha set at 0.01 and adjusted to 0.005 for multiple outcomes, we were unable to reject the null hypothesis that the condition by time interaction effect was outside of the SESOI [-0.1, 0.1]. Thus, from this test we cannot clearly make the claim that there is statistical equivalence between the two interventions in terms of their effects. However, the implications of this study and its results more broadly lend themselves to discussion of important points pertinent to the field.

Despite this representing one of the largest intervention studies to have been conducted in the field of RT research, we still find it difficult to make strong claims regarding the two interventions compared using the stringent inferential standards we set. However, the magnitude of effect estimates and their precision do lead us to further believe that the effects of RT, particularly in trained participants and when comparing interventions, are likely to be far smaller than the typical studies in the field are designed to detect. Most studies in sport and exercise science are considered to be underpowered for even typical moderate-large effects (Abt et al., 2020; Mesquida et al., 2023; Speed & Andersen, 2000), and especially so for RT research given the typical sample sizes used (Steele, Fisher, Smith, et al., 2023) combined with the likely very small effects. Indeed, the main effects for time in both our pre-registered analysis, the exploratory strength analysis, and the additional exploratory analysis for hypertrophy outcomes, all highlight that the effects of RT upon these outcomes, whilst likely real (i.e., interval estimates excluding a zero effect), are small in trained participants, though roughly in line with what can be predicted from prior evidence and theory regarding the effects of RT over time.

The closeness of our a priori predictions with the estimated main effects of time on both hypertrophy and strength are striking. This is particularly so for a field where point predictions and even narrow (or wide) range predictions are almost non-existent. The closeness of the estimates with our predictions lend corroboration to the theory that adaptation in hypertrophy and strength with RT does indeed follow a linear-log function and thus that effects become smaller and smaller with RT experience. Indeed, whilst there may still be real main effects for time in trained persons it may be that these are practically speaking equivalent to no effect, at least after >6 months of prior training experience. The estimates for hypertrophy from our exploratory model for the main effect of time were very close to the predicted effect and within the SESOI for that outcome. Whilst we didn't set a SESOI for strength in this study, the results were also very close to the predicted effect and in a previous work with trained participants, including a large pre-registered study, we found that effects for strength were mostly within the region of practical equivalence set by the population of participants themselves (Carlson et al.,

2022, 2023). Essentially, both hypertrophy and strength appear to be practically maintained in trained people continuing with participation in RT.

This should perhaps further highlight the difficulty for the field in detecting between intervention differences. Where main effects for time are small it is mathematically necessary that, given two interventions with positive simple slopes for time, the interaction effect must be smaller than the main effect. Indeed, as noted for the interventions examined here i.e., fROM and lpROM, prior research suggests that the between condition effects are likely small (Varovic et al., 2024; Wolf et al., 2024). If we allow ourselves here to adopt the lower standards typical of the field we would in fact have to conclude that there is statistical equivalence between the two. In contrast to the convention in the field to set alpha at 0.05 when performing hypothesis testing, we expected effects to be small and, even with setting an equivalence band using a SESOI, our pre-registered analysis was deliberately planned for very high power and thus a low type 2 error rate. As such we lowered our alpha and type 1 error rate to account for this (Lakens, 2022b). Readers of this study might have been disappointed to not have a clear conclusion claimed from our primary pre-registered test; however, readers can interpret out results setting their own alpha similar to the fields typical standards if they wish i.e., 0.05. In this case, adjusting for the two outcomes and using an alpha of 0.025 a reader would conclude that the fROM and lpROM interventions are indeed statistically equivalent within the SESOI of [-0.1, 0.1]. The results of our exploratory hypertrophy analysis removing random slopes for time and utilising all baseline measurements (i.e., t_{-1} and t_0) also strongly support equivalence. We feel that, given the findings presented here, it makes greater sense for researchers in the field of RT, particularly those interested in the effects in trained participants, to examine hypotheses relating to the equivalence of interventions (Mazzolari et al., 2022) as testing for differences even with current inference standards is resource prohibitive[5].

Despite the strengths of this study, we should perhaps pre-empt some of the limitations that might be noted by readers, particularly regarding the outcome operationalisations. It might be suggested that the operationalisations used for hypertrophy (i.e., estimated muscle CSA from circumference and skinfolds) in this study are not gold-standard and thus will have less resolution to detect effects due to greater measurement error. However, we acknowledged this in study planning and explicitly estimated the measurement errors from our prior data using these methods, used this in simulating to determine sample size for the desired statistical power, and also made use of multiple measurements to account for the error associated with single measurements. Indeed, Haun et al. (Haun et al., 2019) have argued that, whilst such measurements might not offer microscopic insights into myofibrillar protein accrual or functional CSA, they are appropriate for examining changes due to RT interventions and particularly so when resource and technical constraints prevent the use of other methods[6]. The

[5]It could also be argued that researchers might instead continue conducting smaller studies but abstain from making strong claims from them. Instead, they should merely reporting the results descriptively and consider that they will one day contribute to a meta-analytic synthesis that might allow stronger claims to be made (Lakens, 2022a).

[6]We also re-analysed data for 67 participants (both untrained and trained) from a previous collection of studies where both circumference and ultrasound based measures of hypertrophy had been used (Gentil et al., 2020). Calculating the pre- to post-intervention SMDs for both methods, and then comparing the methods using a random effects meta-analysis with method as a moderator reveals an SMD for ultrasound measures of 0.34 [95%CI: 0.23, 0.44], for circumference measures of 0.27 [95%CI: 0.19, 0.35], and difference between methods of -0.07 [95%CI: -0.2, 0.06]. Thus, both methods appear clearly able to detect changes in muscle size with relatively similar effect size magnitudes.

same might also be said for our strength operationalization or at least that it does not reflect the measurement of maximal strength as a 1RM would. However, given strength was not our primary outcome, the benefits of high frequency measurement for statistical power and precision outweighed this (Swinton, 2024). Further, we feel confident that out operationalisations reflect the constructs of interest i.e., hypertrophy and strength, given the striking closeness of our estimates to the a priori predictions regarding the effects, predictions which were derived from both theory and empirical evidence which had employed a wide range of operationalisations for both constructs.

Conclusion

This study is perhaps one of the largest to compare the effects of two RT interventions upon hypertrophy; in this case the effects of RT using a fROM or lpROM upon hypertrophy in previously trained participants. We tested statistical equivalence for our primary pre-registered analyses, but we were unable to reject the null hypothesis that the condition by time interaction effect was outside of the SESOI [-0.1, 0.1]. Thus, from this test we cannot clearly make the claim that there is statistical equivalence between the two interventions in terms of their effects. However, both our main pre-registered analysis and additional exploratory analysis suggest that both the main effects of time, and any interaction effects for condition by time, are likely small. The same was the case for strength outcomes. Indeed, the small effects were incredibly similar to what was predicted given the assumption of a theoretical linear-log form to adaptation from RT. These findings are in line with other recent evidence regarding the comparison of fROM and lpROM specifically and do suggest that between condition effects are likely small and practically equivalent. More broadly, this study highlights that the effects of RT in trained persons should be expected to be small and that current studies in the field of RT are woefully underpowered to be able to detect their effects, let alone test between intervention comparisons.

Acknowledgements

We would like to extend our sincere gratitude to all the staff and members at Discover Strength who through their efforts were able to make this study possible.

Funding Information

No direct funding was received for this study. Operational costs for its delivery were absorbed by Discover Strength and so provided *in-kind* value.

Author Contributions

LC, DG, JS, and JF conceived the idea for the project and designed the methods. LC and DG coordinated and oversaw the data collection. JS conducted the statistical analyses and produced data visualisations. JS and JF drafted the initial manuscript. LC, DG, JS, and JF contributed to writing the manuscript. All authors read and approved the final manuscript.

Competing Interests

LC and DG work for Discover Strength. JS and JF provides research consulting for organizations within the health and fitness field.

Data and Supplementary Material Accessibility

All data and code utilised for data preparation and analyses are available in either the Open Science Framework page for this project https://osf.io/t9auy/ or the corresponding GitHub repository https://github.com/jamessteeleii/lenthened_partial_trial. Other supplementary analyses and plots are also available there.

References

Abt, G., Boreham, C., Davison, G., Jackson, R., Nevill, A., Wallace, E., & Williams, M. (2020). Power, precision, and sample size estimation in sport and exercise science research. *Journal of Sports Sciences*, *38*(17), 1933–1935. https://doi.org/10.1080/02640414.2020.1776002

Baechle, T. R., Earle, R. W., & Association (U.S.), N. S. &. C. (2008). *Essentials of Strength Training and Conditioning*. Human Kinetics.

Carlson, L., Gschneidner, D., Steele, J., & Fisher, J. P. (2022). Short-term supervised virtual training maintains intensity of effort and represents an efficacious alternative to traditional studio-based, supervised strength training. *Physiology & Behavior*, *249*, 113748. https://doi.org/10.1016/j.physbeh.2022.113748

Carlson, L., Gschneidner, D., Steele, J., & Fisher, J. P. (2023). The Effects of Training Load During Dietary Intervention Upon Fat Loss: A Randomized Crossover Trial. *Research Quarterly for Exercise and Sport*, *94*(4), 990–1000. https://doi.org/10.1080/02701367.2022.2097625

Gentil, P., Budzynski-Seymour, E., Souza, D., Steele, J., Fisher, J. P., & Bottaro, M. (2020). Evaluating the results of resistance training using ultrasound or flexed arm circumference: A case for keeping it simple? *Journal of Clinical and Translational Research*, *7*(6), 61–65.

Haun, C. T., Vann, C. G., Roberts, B. M., Vigotsky, A. D., Schoenfeld, B. J., & Roberts, M. D. (2019). A Critical Evaluation of the Biological Construct Skeletal Muscle Hypertrophy: Size Matters but So Does the Measurement. *Frontiers in Physiology*, *10*. https://doi.org/10.3389/fphys.2019.00247

Heymsfield, S. B., McManus, C., Smith, J., Stevens, V., & Nixon, D. W. (1982). Anthropometric measurement of muscle mass: Revised equations for calculating bone-free arm muscle area. *The American Journal of Clinical Nutrition*, *36*(4), 680–690. https://doi.org/10.1093/ajcn/36.4.680

Housh, D. J., Housh, T. J., Weir, J. P., Weir, L. L., Johnson, G. O., & Stout, J. R. (1995). Anthropometric estimation of thigh muscle cross-sectional area. *Medicine and Science in Sports and Exercise*, *27*(5), 784–791.

Kassiano, W., Costa, B., Nunes, J. P., Ribeiro, A. S., Schoenfeld, B. J., & Cyrino, E. S. (2023). Which ROMs Lead to Rome? A Systematic Review of the Effects of Range of Motion on Muscle Hypertrophy. *Journal of Strength and Conditioning Research*, *37*(5), 1135–1144. https://doi.org/10.1519/JSC.0000000000004415

Kurz, A. S. (2022). Just use multilevel models for your pre/post RCT data. In *A. Solomon Kurz*.

Lakens, D. (2022a). Sample Size Justification. *Collabra: Psychology*, *8*(1), 33267. https://doi.org/10.1525/collabra.33267

Lakens, D. (2022b). *Improving Your Statistical Inferences*. Zenodo. https://doi.org/10.5281/ZENODO.6409077

Latella, C., van den Hoek, D., Wolf, M., Androulakis-Korakakis, P., Fisher, J. P., & Steele, J. (2024). Using Powerlifting Athletes to Determine Strength Adaptations Across Ages in Males and Females: A Longitudinal Growth Modelling Approach. *Sports Medicine (Auckland, N.Z.)*, *54*(3), 753–774. https://doi.org/10.1007/s40279-023-01962-6

MacInnis, M. J., McGlory, C., Gibala, M. J., & Phillips, S. M. (2017). Investigating human skeletal muscle physiology with unilateral exercise models: When one limb is more powerful than two. *Applied Physiology, Nutrition, and Metabolism = Physiologie Appliquee, Nutrition Et Metabolisme*, *42*(6), 563–570. https://doi.org/10.1139/apnm-2016-0645

Mazzolari, R., Porcelli, S., Bishop, D. J., & Lakens, D. (2022). Myths and methodologies: The use of equivalence and non-inferiority tests for interventional studies in exercise physiology and sport science. *Experimental Physiology*, *107*(3), 201–212. https://doi.org/10.1113/EP090171

Mesquida, C., Murphy, J., Lakens, D., & Warne, J. (2023). Publication bias, statistical power and reporting practices in the Journal of Sports Sciences: Potential barriers to replicability. *Journal of Sports Sciences*, *41*(16), 1507–1517. https://doi.org/10.1080/02640414.2023.2269357

Moreno, E. N., Ayers-Creech, W. A., Gonzalez, S. L., Baxter, H. T., & Buckner, S. L. (2024). Does Performing Resistance Exercise with a Partial Range of Motion at Long Muscle Lengths Maximize Muscle Hypertrophic Adaptations to Training? *Journal of Science in Sport and Exercise*. https://doi.org/10.1007/s42978-024-00301-z

Pallarés, J. G., Hernández-Belmonte, A., Martínez-Cava, A., Vetrovsky, T., Steffl, M., & Courel-Ibáñez, J. (2021). Effects of range of motion on resistance training adaptations: A systematic review and meta-analysis. *Scandinavian Journal of Medicine & Science in Sports*, *31*(10), 1866–1881. https://doi.org/10.1111/sms.14006

Penney, J. (2023). Cautions when normalizing the dependent variable in a regression as a z-score. *Economic Inquiry*, *61*(2), 402–412. https://doi.org/10.1111/ecin.13127

Rodriguez-Sanchez, F., cre, cph, Jackson, C. P., Hutchins, S. D., & Clawson, J. M. (2023). *Grateful: Facilitate Citation of R Packages*.

Schoenfeld, B. J., & Grgic, J. (2020). Effects of range of motion on muscle development during resistance training interventions: A systematic review. *SAGE Open Medicine*, *8*, 2050312120901559. https://doi.org/10.1177/2050312120901559

Schoenfeld, B., Fisher, J., Grgic, J., Haun, C., Helms, E., Phillips, S., Steele, J., & Vigotsky, A. (2021). Resistance Training Recommendations to Maximize Muscle Hypertrophy in an Athletic Population: Position Stand of the IUSCA. *International Journal of Strength and Conditioning*, *1*(1). https://doi.org/10.47206/ijsc.v1i1.81

Speed, H. D., & Andersen, M. B. (2000). What exercise and sport scientists don't understand. *Journal of Science and Medicine in Sport*, *3*(1), 84–92. https://doi.org/10.1016/S1440-2440(00)80051-1

Steele, J., Fisher, J. P., Giessing, J., Androulakis-Korakakis, P., Wolf, M., Kroeske, B., & Reuters, R. (2023). Long-Term Time-Course of Strength Adaptation to Minimal Dose Resistance Training Through Retrospective Longitudinal Growth Modeling. *Research Quarterly for Exercise and Sport*, *94*(4), 913–930. https://doi.org/10.1080/02701367.2022.2070592

Steele, J., Fisher, J. P., Smith, D., Schoenfeld, B. J., Yang, Y., & Nakagawa, S. (2023). Meta-analysis of variation in sport and exercise science: Examples of application within resistance training research. *Journal of Sports Sciences*, *41*(17), 1617–1634. https://doi.org/10.1080/02640414.2023.2286748

Steele, J., Fisher, J., Giessing, J., & Gentil, P. (2017). Clarity in reporting terminology and definitions of set endpoints in resistance training. *Muscle & Nerve*, *56*(3), 368–374. https://doi.org/10.1002/mus.25557

Swinton, P. (2024). *Adequate statistical power in strength and conditioning may be achieved through longer interventions and high frequency outcome measurement.* SportRxiv. https://doi.org/10.51224/SRXIV.364

Varovic, D., Wolf, M., Steele, J., Schoenfeld, B. J., Grgic, J., & Mikulic, P. (2024). *Regional Hypertrophy with Resistance Training—Does Muscle Length Matter? A Systematic Review and Meta-Analysis*. SportRxiv. https://doi.org/10.51224/SRXIV.464

Wolf, M., Androulakis-Korakakis, P., Fisher, J., Schoenfeld, B., & Steele, J. (2023). Partial Vs Full Range of Motion Resistance Training: A Systematic Review and Meta-Analysis. *International Journal of Strength and Conditioning*, *3*(1). https://doi.org/10.47206/ijsc.v3i1.182

Wolf, M., Korakakis, P. A., Piñero, A., Mohan, A. E., Hermann, T., Augustin, F., Sappupo, M., Lin, B., Coleman, M., Burke, R., Nippard, J., Swinton, P. A., & Schoenfeld, B. J. (2024). *Lengthened Partial Repetitions Elicit Similar Muscular Adaptations as a Full Range of Motion During Resistance Training in Trained Individuals*. SportRxiv. https://doi.org/10.51224/SRXIV.455