Statistical Analysis and Results

# 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/c2657/> or the corresponding GitHub repository <https://github.com/jamessteeleii/ROM_regional_hypertrophy>. 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/pgx6v>. As noted, the project was previously pre-registered however in hindsight we realise that the details of our analysis plan were imprecise and left open many researcher degrees of freedom. Thus, we present the planned analyses as closely as possible given the pre-registration as written and our original intention, but note where we have deviated from this plan below. Further, given the ambiguity we have conducted several additional analyses, including varying the priors used and the model parametrization, the methods and results of which are described in full detail in the supplementary materials here <https://osf.io/rqavs>. In the main text here we report only the pre-registered main models.

All analyses have been conducted within a Bayesian meta-analytic framework and all posterior estimates and their precision, along with conclusions based upon them, will be interpreted continuously and probabilistically, considering priors, data quality, and all within the context of each outcome and the assumptions of the model employed as the estimator (Kruschke & Liddell, 2018). We deviate from the pre-registration in the number of sampling iterations (pre-registered as 6000) used as we include comparisons between all models fit in the supplementary materials using Bayes Factors and the Savage-Dickey ratio where it is recommended that at least 40000 iterations are used to obtain precise Bayes-Factors (Gronau et al., 2020). Trace plots were produced along with values to examine whether chains had converged, and posterior predictive checks for each model were also examined to understand the model implied distributions.

## Effect sizes

We explored effects calculated for within arm pre- to post-intervention (and for studies with multiple post baseline time points pre- to each time point) as the standardized mean change using raw score standardization with heteroscedastic population variances (SMD) (Bonett, 2008) given it is known that variances scale with mean values in RT study outcomes (Steele et al., 2023) and so a pre-post intervention effect upon the mean will influence this. We also examined the log transformed response ratio (lnRR) (Lajeunesse, 2011), which was exponentiated back to the percentage change scale after model fitting (though note that all prior distributions were set on the lnRR scale directly) accounting for the total variance in the model when doing so for the meta-analytic predicted effects (Nakagawa et al., 2017; Spake et al., 2023), as this effect size statistic is unaffected (except in its sampling variance) by the estimates for standard deviations within individual studies which are likely underpowered in the typical sample sizes found in the RT literature (Steele et al., 2023). The use of both additive and multiplicative effect sizes also allows us to explore the sensitivity of interaction effects to scaling as interactions, our primary estimand of interest, are very sensitive to this (Rohrer & Arslan, 2021; Spake et al., 2023). Effects were weighted in each model by their inverse sampling variance.

## Models

The primary estimand of interest was the population level (i.e, fixed effect) muscle length by site of measurement interaction. In each model, as per the preregistration, muscle length and site of measurement were centred at 50%, and rescaled to be on the interval[[1]](#footnote-25). As such, the population level coefficients in each of the models corresponded to the overall average effect of RT on hypertrophy when at a muscle length of 50% at a site of measurement of 50% (i.e., the intercept: in each model below), the slope of the difference i.e., comparison between 0% and 100% muscle length at a site of measurement of 50% (i.e., muscle length coefficient: in each model below), the slope of the difference i.e., comparison between 0% and 100% site of measurement at a muscle length of 50% (i.e., site of measurement coefficient: in each model below), and the slope of the difference i.e., comparison between 0% and 100% site of measurement on the slope of the difference i.e., comparison between 0% and 100% muscle length (i.e., muscle length by site of measurement interaction coefficient: in each model below). Notably, the interpretation of continuous by continuous predictors can be quite challenging. As such, we present for each of these models draws from the posterior of the expectation of the predicted global grand mean across muscle length and at three levels of site of measurement (25%, 50%, and 75%) which shows the predicted effect size magnitudes at particular combinations of muscle length and site of measurement, in addition to the slopes for muscle length at three levels of site of measurement (25%, 50%, and 75%). The slopes for muscle length were transformed to reflect the average contrast in mean muscle length between “shorter” and “longer” conditions in the included studies which was . Thus the slopes reflected a difference in muscle length of 21.8% e.g., the slope of the difference between 32.4% (the average “short” condition in the included studies) and 54.2% muscle length which shows the magnitude of the difference in effect size for a 21.8% difference in muscle length at different sites. We present the predicted values and slopes as mean and 95% quantile intervals. We also, whilst not pre-registered for this project, agreed upon a smallest effect size of interest on both the standardized mean change [[2]](#footnote-26) and the percentage change scales and thus set these as regions of practical equivalence (ROPE). This allows us to also examine the probability that the slopes for muscle length might produce a meaningful effect (i.e., greater than the smallest effect size of interest) by examining the mass of the posterior distribution exceeding the upper limits of the ROPE, and also the percentage of the posterior distributions mass that was within the ROPE thus reflecting the probability of practically equivalent effects.

For the secondary predictor models we explored the muscle length by site of measurement by each additional predictor (e.g., upper or lower body OR muscle group OR muscle action) interaction respectively. For these models, we only present the predicted effect size magnitudes at particular combinations of muscle length and site of measurement similarly to the above.

### Pre-Registered Main Model

As noted, the pre-registered main model involved population level effects for the intercept, slope of muscle length, slope of measurement site, and the muscle length by measurement site interaction. The model also included random intercepts for study, arm, and effect levels. The model equation was as follows:

where is the th effect size (), here the SMD or lnRR, from the th arm () for the th study (), and , , and are the random intercepts for study, arm, and effect respectively. Prior distributions[[3]](#footnote-29) taken from Wolf et al. (2023) for the SMD model were (note, values rounded; plots for the population level effect distributions can be seen in the supplementary materials here <https://osf.io/uxhdj>):

Prior distributions taken from Wolf et al. (2023) for the lnRR model were (note, values rounded; plots for the population level effect distributions can be seen in the supplementary materials here <https://osf.io/tvpes>):

### Secondary Predictor Models - Uninformed Priors

Although we noted these in the pre-registration we do not focus on them in the present manuscript instead focusing on the primary estimand noted above of the muscle length by measurement site interaction. We treat these secondary predictor models as highly exploratory given the amount of data available and the corresponding uncertainty of inferences, and present them only in the supplementary materials (see <https://osf.io/tgzpk>, <https://osf.io/f86ng>, and <https://osf.io/gp2vr> for the upper or lower body, muscle group, and muscle action SMD models respectively and <https://osf.io/hxbv6>, <https://osf.io/w8mbg>, and <https://osf.io/9mhcu> for the upper or lower body, muscle group, and muscle action lnRR models respectively). In addition, and not pre-registered either, we included a model comparing studies which manipulated mean muscle length by means of range of motion manipulation, or by means of exercise selection (see <https://osf.io/pbqwe> and <https://osf.io/9snkh>). For reference these models were the same parametrization as the pre-registered model using the same priors as noted above, with the exception of the additional categorical predictor of either upper or lower body OR muscle group OR muscle action added under a deviation coding scheme (i.e., such that the coefficients for each level were in comparison to the overall mean reflected by the intercept). The added predictors used default uninformative priors of .

# Results

The final models presented all included 184 effects nested within 22 intervention arms extracted from 12 studies.

## Pre-Registered Main Model

For the main pre-registered model utilizing priors from Wolf et al. (2023) the predicted effect size magnitudes across muscle length and at three levels of site of measurement (25%, 50%, and 75%) in addition to the slopes for muscle length (transformed to be the slope of a difference in muscle length of 21.8% e.g., the slope of the difference between ~32.4% and ~54.2% muscle length) at three levels of site of measurement (25%, 50%, and 75%) can be seen in [Figure 1](#fig-wolf-SMD-plot) for the SMD model, and [Figure 2](#fig-wolf-lnRR-plot) for the lnRR model.

For the SMD model the magnitude of muscle length slope was 0.04 [95% quantile interval: -0.07, 0.15] at the 25% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 16.54% and percentage within the ROPE (i.e., -0.1,0.1) of 82.7%, 0.07 [95% quantile interval: -0.02,0.15] at the 50% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 22.3% and percentage within the ROPE (i.e., -0.1,0.1) of 77.68%, and 0.09 [95% quantile interval: -0.01,0.19] at the 75% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 40.96% and percentage within the ROPE (i.e., -0.1,0.1) of 59.02%. There was considerable heterogeneity of effects relative to the magnitude of the population level effects, particularly at the study level, with 0.04 [95% quantile interval: 0.03,0.35], 0.03 [95% quantile interval: 0.01,0.26], and 0.01 [95% quantile interval: 0,0.07].

For the lnRR model the magnitude of muscle length slope was 0.48% [95% quantile interval: -1.99%, 3.13%] at the 25% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 3.09% and percentage within the ROPE (i.e., -0.1,0.1) of 96.66%, 1.14% [95% quantile interval: -0.84%,3.13%] at the 50% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 3.38% and percentage within the ROPE (i.e., -0.1,0.1) of 96.62%, and 1.8% [95% quantile interval: -0.52%,4.26%] at the 75% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 16.55% and percentage within the ROPE (i.e., -0.1,0.1) of 83.45%. There was considerable heterogeneity of effects relative to the magnitude of the population level effects, particularly at the study level, with 1.99% [95% quantile interval: 3.6%,14.63%], 1.13% [95% quantile interval: 1.71%,7.96%], and 0.45% [95% quantile interval: 0.37%,1.67%].

Model diagnostics can be seen in the supplementary materials here: <https://osf.io/3ybcs>.

|  |
| --- |
| Figure 1: Results from primary pre-registered main model for standardized mean difference effects. |

|  |
| --- |
| Figure 2: Results from primary pre-registered main model for exponentiated log response ratio effects. |

# References

Bonett, D. G. (2008). Confidence intervals for standardized linear contrasts of means. *Psychological Methods*, *13*(2), 99–109. <https://doi.org/10.1037/1082-989X.13.2.99>

Gronau, Q. F., Singmann, H., & Wagenmakers, E.-J. (2020). Bridgesampling: An R Package for Estimating Normalizing Constants. *Journal of Statistical Software*, *92*, 1–29. <https://doi.org/10.18637/jss.v092.i10>

Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review*, *25*(1), 178–206. <https://doi.org/10.3758/s13423-016-1221-4>

Lajeunesse, M. J. (2011). On the meta-analysis of response ratios for studies with correlated and multi-group designs. *Ecology*, *92*(11), 2049–2055. <https://doi.org/10.1890/11-0423.1>

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of The Royal Society Interface*, *14*(134), 20170213. <https://doi.org/10.1098/rsif.2017.0213>

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

Rohrer, J. M., & Arslan, R. C. (2021). Precise Answers to Vague Questions: Issues With Interactions. *Advances in Methods and Practices in Psychological Science*, *4*(2), 25152459211007368. <https://doi.org/10.1177/25152459211007368>

Spake, R., Bowler, D. E., Callaghan, C. T., Blowes, S. A., Doncaster, C. P., Antão, L. H., Nakagawa, S., McElreath, R., & Chase, J. M. (2023). Understanding “it depends” in ecology: A guide to hypothesising, visualising and interpreting statistical interactions. *Biological Reviews*, *98*(4), 983–1002. <https://doi.org/10.1111/brv.12939>

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>

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>

1. The rescaling was in order to have the values for coefficients in the model on a similar scale as the intercept values so that when setting initial values for Monte Carlo Markov Chain sampling these could be set to similar values. [↑](#footnote-ref-25)
2. Though some of the authors here were those who had been involved in eliciting the smallest effect size of interest used for a currently pre-registered trial examining the effects lengthened partial training upon hypertrophy (see <https://osf.io/9sgjk>). [↑](#footnote-ref-26)
3. We obtained priors from studies in the dataset of Wolf et al. (2023) by firstly calculating the within arm pre- to post-intervention changes for the partial range of motion groups only (both SMD and lnRR effect sizes as detailed above) for only muscle size outcomes. We excluded any studies that were already included in the present dataset. Muscle length categorised as short or long was then recoded to be -0.5 and 0.5 respectively such that the predictor was centred; we assumed that the typical difference between short and long was similar to the ~21.8% as noted above in extracting slopes for reporting and so these codes corresponded to ~32.4% and ~54.2% muscle length respectively. Measurement site was also centred at 50%. We then fit a model with the same parametrization as the present pre-registered model in [Equation 1](#eq-main-model) with the primary difference being that muscle length was categorical and the coefficient in the model fit to the Wolf et al. (2023) data reflected the slope of the difference i.e., comparison between ~32.4% and ~54.2% muscle length. We set weakly regularising priors of for the population parameters for this model as with default uninformative priors chains did not converge. Other priors were left as defaults. We then extracted, assuming distributions, the hyperparameters , , and for the following parameters: intercept, muscle length coefficient, measurement site coefficient, muscle length by measurement site interaction, and the random effects standard deviations for the study, arm, and effect level intercepts. [↑](#footnote-ref-29)