

Statistical Analysis and Results

All models and model comparisons

This supplementary document contains the full statistical methods and results for all models, both the pre-registered main model and additional supplementary models varying priors and also model parametrisations, for the project titled “*The effects of muscle length during resistance training on regional hypertrophy adaptations: A systematic review and meta-analysis*” by Dorian Varovic, Milo Wolf, Brad Jon Schoenfeld, James Steele, and Pavle Mikulić.

Statistical Analysis

All code utilised 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/jamessteelei/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.

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 model comparisons with Bayes Factors using 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 \hat{R} values to examine whether chains had converged, and posterior predictive checks for each model were also examined to understand the model implied distributions.

We preregistered that we would obtain informed priors for relevant model parameters from the previous meta-analysis of Wolf et al. (2023) excluding any studies included in their posterior

estimates that would also be included in the likelihood of the models in the present analyses. We did not however detail exactly how we would obtain these priors and so provide details below. In addition to the pre-registered main models using these priors, we also present a different parametrisation of the main model (employing random slopes as noted below) as well as two models utilising different informed priors. The reason for the two additional sets of informative priors was due to disagreement between the author group (where one coauthor, JS, held far more skeptical priors based upon other background evidence not contained in the likelihood) and that the priors taken from Wolf et al. (2023) were, relatively speaking, uninformative due to their imprecision and had little impact on these separate priors. Whilst we did not preregister these, exploring the sensitivity of posterior inferences to the priors used is a common part of Bayesian workflows (Dragicevic et al., 2019; Kallioinen et al., 2023). As such we present them here, in addition to a completely uninformed default prior model (we also employed uninformative default priors for the secondary predictor models), to demonstrate how the evidence from the studies included regarding the effects of muscle length, site of measurement, and their interaction, might shift the posterior beliefs of those coming from different prior beliefs about their effects.

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 resistance training 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 resistance training 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

In the main pre-registered models, and those secondary models including addition of random slopes and different informed priors, 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 $(-0.5, 0.5)$ interval¹. As such, the population level coefficients in each of the models corresponded to the overall average effect of resistance training on hypertrophy when at a muscle length of 50% at a site of measurement of 50% (i.e., the intercept: μ_{α_i} 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: β_1 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: β_2 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: β_3 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 (transformed to be the slope of a difference in muscle length of 50% e.g., the slope of the difference between 25% and 75% muscle length) at three levels of site of measurement (25%, 50%, and 75%) which shows the magnitude of the difference in effect size for a 50% 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 standardised mean change $(-0.1, 0.1)$ ² and the percentage change scales $(-3\%, 3\%)$ 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.

¹The 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.

²Though 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>).

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:

$$\begin{aligned}
 \hat{\mu}_{ijk} &\sim N(\mu, \sigma_{ijk}) \\
 \mu &= \alpha_i + \alpha_j + \alpha_k + \beta_1(\text{muscle length}_{centred}) + \beta_2(\text{site}_{centred}) + \beta_3(\text{muscle length}_{centred} \times \text{site}_{centred}) \\
 \alpha_i &\sim N(\mu_{\alpha_i}, \sigma_{\alpha_i}), \text{ for study } i = 1, \dots, I \\
 \alpha_j &\sim N(0, \sigma_{\alpha_j}), \text{ for arm } j = 1, \dots, J \\
 \alpha_k &\sim N(0, \sigma_{\alpha_k}), \text{ for effect } k = 1, \dots, K
 \end{aligned} \tag{1}$$

where $\hat{\mu}_{ijk}$ is the k th effect size ($k = 1, \dots, K$), here the SMD or lnRR, from the j th arm ($j = 1, \dots, J$) for the i th study ($i = 1, \dots, I$), and α_i , α_j , and α_k are the random intercepts for study, arm, and effect respectively. Prior distributions³ 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>):

³We 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 ~50% as noted above in extracting slopes for reporting in other models and so these codes corresponded to ~25% and ~75% muscle length respectively. Measurement site was also centred at 50%. We then fit a model with the same parametrisation as the present pre-registered model in Equation 1 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 ~25% and ~75% muscle length. We set weakly regularising priors of *student t* ($df = 3, \mu = 0, \sigma = 1$) 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 *student t* distributions, the hyperparameters df , μ , and σ 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.

$$\begin{aligned}
\mu_{\alpha_i} &\sim \text{student } t(df = 4.17, \mu = 0.16, \sigma = 0.53) \\
\sigma_{\alpha_i} &\sim \text{half student } t(df = 2.32, \mu = 0.79, \sigma = 0.60) \\
\sigma_{\alpha_j} &\sim \text{half student } t(df = 2.32, \mu = 0.62, \sigma = 0.46) \\
\sigma_{\alpha_k} &\sim \text{half student } t(df = 3.53, \mu = 0.14, \sigma = 0.10) \\
\beta_1 &\sim \text{student } t(df = 4.39, \mu = 0.40, \sigma = 1.15) \\
\beta_2 &\sim \text{student } t(df = 7.03, \mu = -0.02, \sigma = 0.91) \\
\beta_3 &\sim \text{student } t(df = 4.45, \mu = -0.09, \sigma = 1.95)
\end{aligned} \tag{2}$$

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>):

$$\begin{aligned}
\mu_{\alpha_i} &\sim \text{student } t(df = 2.82, \mu = 0.04, \sigma = 0.39) \\
\sigma_{\alpha_i} &\sim \text{half student } t(df = 1.92, \mu = 0.55, \sigma = 0.47) \\
\sigma_{\alpha_j} &\sim \text{half student } t(df = 1.96, \mu = 0.43, \sigma = 0.35) \\
\sigma_{\alpha_k} &\sim \text{half student } t(df = 7.38, \mu = 0.06, \sigma = 0.04) \\
\beta_1 &\sim \text{student } t(df = 2.80, \mu = 0.10, \sigma = 0.80) \\
\beta_2 &\sim \text{student } t(df = 7.74, \mu = 0.10, \sigma = 0.48) \\
\beta_3 &\sim \text{student } t(df = 5.95, \mu = 0.08, \sigma = 1.70)
\end{aligned} \tag{3}$$

Additional Models

Main Model - Uninformed Priors

We include a model following the same parametrisation as the pre-registered model (Equation 1) above (i.e., population level effects for the intercept, slope of muscle length, slope of measurement site, and the muscle length by measurement site interaction and also random intercepts for study, arm, and effect levels) but with uninformed default priors purely for comparative purposes. The uninformative default priors were as follows:

$$\begin{aligned}
\mu_{\alpha_i} &\sim \text{uniform}(lb = -\infty, ub = \infty) \\
\sigma_{\alpha_i} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 2.5) \\
\sigma_{\alpha_j} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 2.5) \\
\sigma_{\alpha_k} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 2.5) \\
\beta_1 &\sim \text{uniform}(lb = -\infty, ub = \infty) \\
\beta_2 &\sim \text{uniform}(lb = -\infty, ub = \infty) \\
\beta_3 &\sim \text{uniform}(lb = -\infty, ub = \infty)
\end{aligned} \tag{4}$$

Main Model - Addition of Random Slopes

In hindsight we realised when we came to analyse the data that we probably should have included in our pre-registration the inclusion of maximal random effects at the study level i.e., random slopes for both muscle length and measurement site, to reflect the assumption that the relationships between these variables and effect sizes might vary across studies. As such, we include a model following the same parametrisation as the pre-registered model (i.e., population level effects for the intercept, slope of muscle length, slope of measurement site, and the muscle length by measurement site interaction and also random intercepts for study, arm, and effect levels; Equation 1) but with the *addition* of random slopes for both muscle length and measurement site at the study level for comparative purposes. The model equation was as follows:

$$\begin{aligned}
\hat{\mu}_{ijk} &\sim N(\mu, \sigma_{ijk}) \\
\mu &= \alpha_i + \alpha_j + \alpha_k + \beta_1(\text{muscle length}_{\text{centred}}) + \beta_2(\text{site}_{\text{centred}}) + \beta_3(\text{muscle length}_{\text{centred}} \times \text{site}_{\text{centred}}) \\
\alpha_j &\sim N(0, \sigma_{\alpha_j}), \text{ for arm } j = 1, \dots, J \\
\alpha_k &\sim N(0, \sigma_{\alpha_k}), \text{ for effect } k = 1, \dots, K \\
\begin{pmatrix} \alpha_i \\ \beta_{1i} \\ \beta_{2i} \end{pmatrix} &\sim N \left(\begin{pmatrix} \mu_{\alpha_i} \\ \mu_{\beta_{1i}} \\ \mu_{\beta_{2i}} \end{pmatrix}, \begin{pmatrix} \sigma_{\alpha_i} & 0 & 0 \\ 0 & \sigma_{\beta_{1i}} & 0 \\ 0 & 0 & \sigma_{\beta_{2i}} \end{pmatrix} R \begin{pmatrix} \sigma_{\alpha_i} & 0 & 0 \\ 0 & \sigma_{\beta_{1i}} & 0 \\ 0 & 0 & \sigma_{\beta_{2i}} \end{pmatrix} \right), \text{ for study } i = 1, \dots, I
\end{aligned} \tag{5}$$

where $\hat{\mu}_{ijk}$ is the k th effect size ($k = 1, \dots, K$), here the SMD or lnRR, from the j th arm ($j = 1, \dots, J$) for the i th study ($i = 1, \dots, I$), α_i , α_j , and α_k are the random intercepts for study, arm, and effect respectively, and β_{1i} and β_{2i} were the random slopes within study for muscle length and measurement site respectively. The priors were the same as those used for the pre-registered model (Equation 2 and Equation 3) with the addition of a *student* $t(df = 3, \mu = 0, \sigma = 2.5)$ prior on both muscle length and measurement site random slopes, and a $LKJcorr(1)$ prior on the covariance matrix R .

Main Model with Random Slopes - J. Steele Priors

Though not included in the pre-registration one of the authors (JS) held a different set of priors for various model parameters based upon a previous large meta-analysis they conducted of randomised trials of resistance training interventions in general compared to non-training controls (Steele et al., 2023). Given the scale of this meta-analysis involving over 100 studies the overall estimate for the general effects of resistance training upon hypertrophy was quite precise (SMD = 0.34 95%CI: 0.29, 0.39; exponentiated lnRR = 5.13% 95%CI: 4.08%, 6.18%). Further, it provided an estimate of the between unit variance in effect sizes at the study, arm, and effect levels (other variance parameters such as the random slopes employed default weakly regularising priors). JS felt that, given the magnitude of the estimate from Steele et al. (2023) any interaction effect would necessarily need to be far smaller; in this case, the effect of muscle length and the effect of measurement site in the present analysis reflects an interaction effect on the general effect of resistance training i.e., the difference in intervention effects (i.e., the counterfactual comparison of a resistance training intervention with a non-training control) between two different conditions such as short vs long muscle length resistance training interventions or proximal vs distal measurement sites. If there are any positive effects of training at longer muscle lengths as posited by others (Kassiano et al., 2023) then they necessarily had to be small, though it was unlikely that any effect was to be negative. In addition the effects of measurement site would have to be small, though it seemed less clear that there would necessarily be effects favouring more proximal or distal sites (Zabaleta-Korta et al., 2020). Lastly, the interaction of muscle length and measurement site would again necessarily have to be smaller given that in the broader conceptualisation these reflect a kind of three-way interaction; in essence the difference in intervention effects (i.e., the counterfactual comparison of a resistance training intervention with a non-training control) between short vs long muscle length resistance training interventions at different measurement sites. Given the precision of the overall estimates it seemed reasonable to set a skew normal prior with mass mostly towards null effects but permitting small effects for the muscle length coefficient, and a student t distribution centred on null effects but permitting small effects in either direction, in the approximate range of the confidence interval half-widths of the estimates from Steele et al. (2023). The muscle length by measurement site interaction effect was then also a student t distribution centred at null with half the scale of the coefficient for measurement site. Notably, these priors were far stronger and more skeptical of large effects than the priors taken from the Wolf et al. (2023) model⁴. Note that priors for the muscle length, measurement site, and interaction, were set considering what might be reasonable given prior knowledge for a difference of 50% muscle length or measurement site (e.g., 25% to 75%) to reflect a “typical” short vs long or proximal vs distal comparison.

Thus a model was fit to enable us to draw posterior inferences given the strong skeptical priors

⁴Note, the model used to extract priors from Wolf et al. (2023) described in footnote 3 was also fit using the priors elicited from JS which revealed that the data used from Wolf et al. (2023) had little influence on the posterior estimates and conclusions that JS would have held prior to the present analyses (see prior and posterior distributions in the supplemental material here for SMD <https://osf.io/h6ks7> and lnRR <https://osf.io/w7zmk>).

elicited from JS. The model parametrisation was the same as that used above (Equation 5) where random slopes were added to the main pre-registered model parametrisation (i.e., population level effects for the intercept, slope of muscle length, slope of measurement site, and the muscle length by measurement site interaction, random intercepts for study, arm, and effect levels, and random slopes for both muscle length and measurement site at the study level). The priors were as follows for the SMD model (note, values rounded; plots for the population level effect distributions can be seen in the supplementary materials here <https://osf.io/g8stn>):

$$\begin{aligned}
\mu_{\alpha_i} &\sim \text{student } t(df = 103, \mu = 0.34, \sigma = 0.03) \\
\sigma_{\alpha_i} &\sim \text{half student } t(df = 3, \mu = 0.21, \sigma = 0.05) \\
\sigma_{\alpha_j} &\sim \text{half student } t(df = 3, \mu = 0.02, \sigma = 0.05) \\
\sigma_{\alpha_k} &\sim \text{half student } t(df = 3, \mu = 0.05, \sigma = 0.04) \\
\sigma_{\beta_{1i}} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 2.5) \\
\sigma_{\beta_{2i}} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 2.5) \\
\beta_1 &\sim \text{skew normal } (\mu = 0, \sigma = 0.1, \alpha = 5) \\
\beta_2 &\sim \text{student } t(df = 3, \mu = 0, \sigma = 0.04) \\
\beta_3 &\sim \text{student } t(df = 3, \mu = 0, \sigma = 0.02) \\
R &\sim LKJcorr(1)
\end{aligned} \tag{6}$$

The priors were as follows for the lnRR model (note, values rounded; plots for the population level effect distributions can be seen in the supplementary materials here <https://osf.io/4fhvz>):

$$\begin{aligned}
\mu_{\alpha_i} &\sim \text{student } t(df = 103, \mu = 0.05, \sigma = 0.005) \\
\sigma_{\alpha_i} &\sim \text{half student } t(df = 3, \mu = 0.02, \sigma = 0.005) \\
\sigma_{\alpha_j} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 0.006) \\
\sigma_{\alpha_k} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 0.006) \\
\sigma_{\beta_{1i}} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 2.5) \\
\sigma_{\beta_{2i}} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 2.5) \\
\beta_1 &\sim \text{skew normal } (\mu = 0, \sigma = 0.016, \alpha = 5) \\
\beta_2 &\sim \text{student } t(df = 3, \mu = 0, \sigma = 0.008) \\
\beta_3 &\sim \text{student } t(df = 3, \mu = 0, \sigma = 0.004) \\
R &\sim LKJcorr(1)
\end{aligned} \tag{7}$$

Main Model with Random Slopes - Other Authors Priors

After JS raised his skeptical priors with the other authors they too felt it might be appropriate to consider what priors they might have held absent the Wolf et al. (2023) data. The other authors agreed with JS in the priors for the overall intercept and random intercept variances as taken from Steele et al. (2023), though they did not express the same degree of skepticism as JS in the prior effects for muscle length, measurement site, and their interaction instead wanting to permit larger modal positive effects and a wider range of possible effects for each parameter.

Thus a model was fit to enable us to draw posterior inferences given the less skeptical priors elicited from the other authors. The model parametrisation was the same as that used above (Equation 5) where random slopes were added to the main pre-registered model parametrisation (i.e., population level effects for the intercept, slope of muscle length, slope of measurement site, and the muscle length by measurement site interaction, random intercepts for study, arm, and effect levels, and random slopes for both muscle length and measurement site at the study level). The priors were as follows for the SMD model (note, values rounded; plots for the population level effect distributions can be seen in the supplementary materials here <https://osf.io/rhs7d>):

$$\begin{aligned}
\mu_{\alpha_i} &\sim \text{student } t(df = 103, \mu = 0.34, \sigma = 0.03) \\
\sigma_{\alpha_i} &\sim \text{half student } t(df = 3, \mu = 0.21, \sigma = 0.05) \\
\sigma_{\alpha_j} &\sim \text{half student } t(df = 3, \mu = 0.02, \sigma = 0.05) \\
\sigma_{\alpha_k} &\sim \text{half student } t(df = 3, \mu = 0.05, \sigma = 0.04) \\
\sigma_{\beta_{1i}} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 2.5) \\
\sigma_{\beta_{2i}} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 2.5) \\
\beta_1 &\sim \text{student } t(df = 3, \mu = 0.4, \sigma = 0.2) \\
\beta_2 &\sim \text{student } t(df = 3, \mu = 0.4, \sigma = 0.2) \\
\beta_3 &\sim \text{student } t(df = 3, \mu = 0.2, \sigma = 0.1) \\
R &\sim LKJcorr(1)
\end{aligned} \tag{8}$$

The priors were as follows for the lnRR model (note, values rounded; plots for the population level effect distributions can be seen in the supplementary materials here <https://osf.io/e5hz3>):

$$\begin{aligned}
\mu_{\alpha_i} &\sim \text{student } t(df = 103, \mu = 0.05, \sigma = 0.005) \\
\sigma_{\alpha_i} &\sim \text{half student } t(df = 3, \mu = 0.02, \sigma = 0.005) \\
\sigma_{\alpha_j} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 0.006) \\
\sigma_{\alpha_k} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 0.006) \\
\sigma_{\beta_{1i}} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 2.5) \\
\sigma_{\beta_{2i}} &\sim \text{half student } t(df = 3, \mu = 0, \sigma = 2.5) \\
\beta_1 &\sim \text{student } t(df = 3, \mu = 0.095, \sigma = 0.04) \\
\beta_2 &\sim \text{student } t(df = 3, \mu = 0.095, \sigma = 0.04) \\
\beta_3 &\sim \text{student } t(df = 3, \mu = 0.048, \sigma = 0.02) \\
R &\sim LKJcorr(1)
\end{aligned} \tag{9}$$

Model Comparisons

Given the addition of several models to the present analysis varying both the parametrisation and the prior distributions employed, and that as we acknowledge the pre-registered analysis may not have been sufficiently detailed to constrain researcher degrees of freedom nor did it take into account assumptions we should have made about models and priors, we felt it prudent to compare these models. Thus we calculated Bayes Factors quantifying the ratio:

$$BF_{12} = \frac{P(Data|Model_1)}{P(Data|Model_2)} \tag{10}$$

This answers the question “under which model are the observed data more probable?” where positive values indicate support for the numerator (i.e., $Model_1$) and negative values indicate support for the denominator (i.e., $Model_2$). The Bayes Factors were transformed as $2 \times \ln(BF)$ and then compared to Kass and Raftery’s (1995) scale regarding evidence favouring $Model_1$ vs $Model_2$ (i.e., 0 to 2 = weak, 2 to 6 = positive, 6 to 10 = strong, 10 or greater = very strong).

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). For reference these models were the same parametrisation 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 $uniform(lb = -\infty, ub = \infty)$.

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 utilising 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 50% e.g., the slope of the difference between 25% and 75% muscle length) at three levels of site of measurement (25%, 50%, and 75%) can be seen in Figure 1 for the SMD model, and Figure 2 for the lnRR model.

For the SMD model the magnitude of muscle length slope was 0.12 [95% quantile interval: -0.15, 0.37] at the 25% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 55.12% and percentage within the ROPE (i.e., -0.1,0.1) of 39.18%, 0.17 [95% quantile interval: -0.04,0.37] at the 50% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 75.32% and percentage within the ROPE (i.e., -0.1,0.1) of 24%, and 0.23 [95% quantile interval: -0.01,0.46] at the 75% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 85.05% and percentage within the ROPE (i.e., -0.1,0.1) of 14.54%. There was considerable heterogeneity of effects relative to the magnitude of the population level effects, particularly at the study level, with $\tau_{study} = 0.09$ [95% quantile interval: 0.03,0.36], $\tau_{arm} = 0.06$ [95% quantile interval: 0.01,0.25], and $\tau_{effect} = 0.01$ [95% quantile interval: 0,0.07].

For the lnRR model the magnitude of muscle length slope was 1.56% [95% quantile interval: -4.33%, 7.87%] at the 25% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 32.37% and percentage within the ROPE (i.e., -0.1,0.1) of 60.98%, 3.17% [95% quantile interval: -1.72%,8.08%] at the 50% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 52.73% and percentage within the ROPE (i.e., -0.1,0.1) of 46.63%, and 4.8% [95% quantile interval: -0.85%,10.78%] at the 75% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 72.96% and percentage within the ROPE (i.e., -0.1,0.1) of 26.74%. There was considerable heterogeneity of effects relative to the magnitude of

the population level effects, particularly at the study level, with $\tau_{study} = 4.26\%$ [95% quantile interval: 3.88%,14.88%], $\tau_{arm} = 2.01\%$ [95% quantile interval: 1.28%,7.64%], and $\tau_{effect} = 0.6\%$ [95% quantile interval: 0.37%,1.69%].

Model diagnostics can be seen in the supplementary materials here: <https://osf.io/3ybc5>.

Interaction between mean muscle length and site of measurement

Global grand mean and 95% quantile intervals presented for predictions and slopes at 25%, 50%, and 75% of centred site of measurement

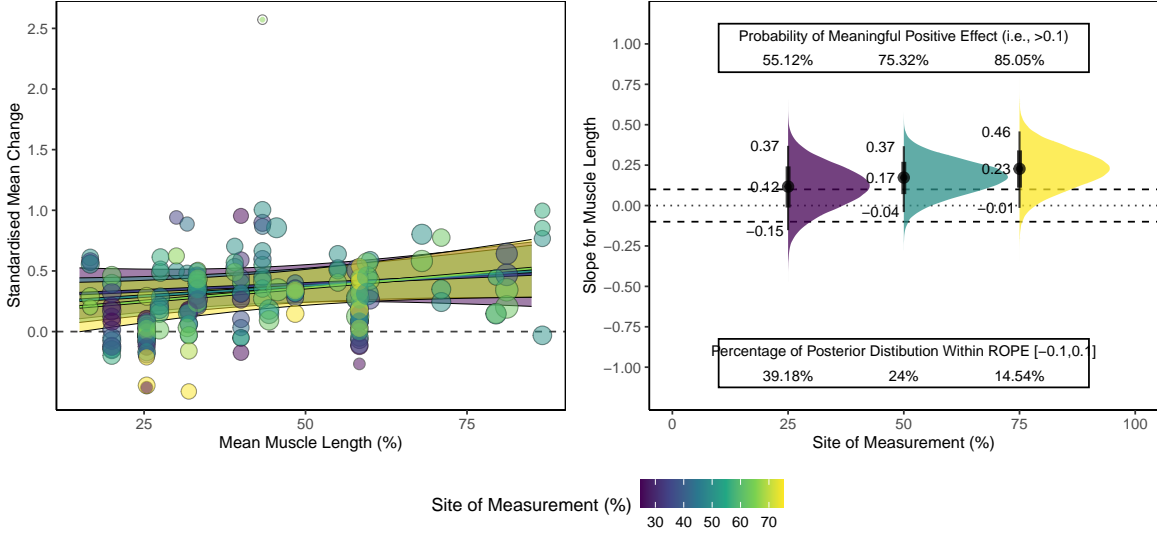


Figure 1: Results from primary pre-registered main model for standardised mean difference effects.

Additional Models

Main Model - Uninformed Priors

For the main model parametrisation yet with uninformative default priors 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 50% e.g., the slope of the difference between 25% and 75% muscle length) at three levels of site of measurement (25%, 50%, and 75%) can be seen in Figure 3 for the SMD model, and Figure 4 for the lnRR model.

For the SMD model the magnitude of muscle length slope was 0.11 [95% quantile interval: -0.17, 0.37] at the 25% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 53.66% and percentage within the ROPE (i.e., -0.1,0.1) of 39.62%, 0.17 [95% quantile interval: -0.04,0.37] at the 50% measurement site with probability of a meaningful positive effect (i.e., 0.1)

Interaction between mean muscle length and site of measurement

Global grand mean and 95% quantile intervals presented for predictions and slopes at 25%, 50%, and 75% of centred site of measurement

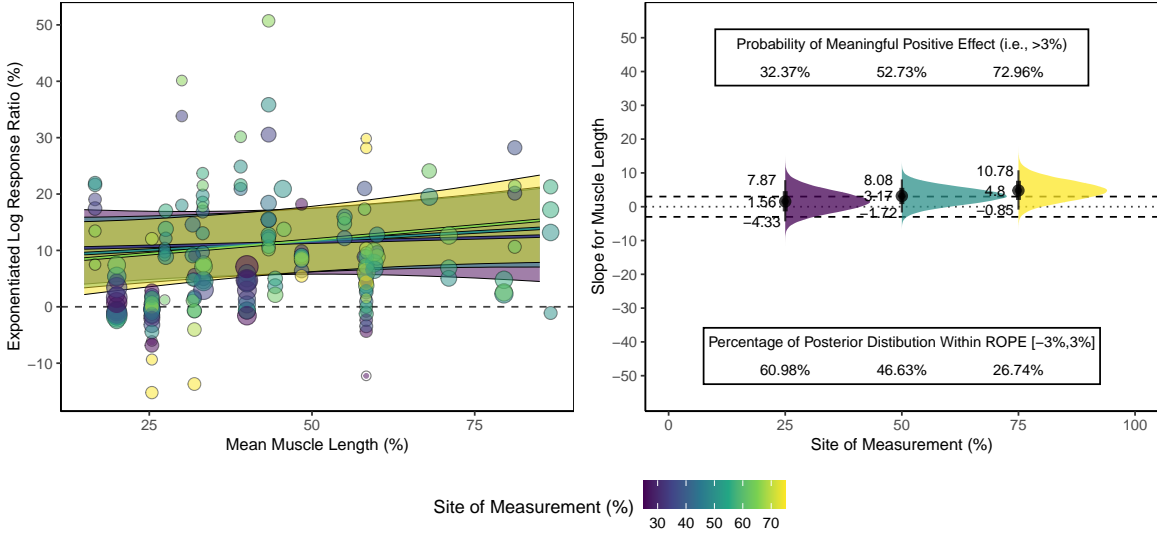


Figure 2: Results from primary pre-registered main model for exponentiated log response ratio effects.

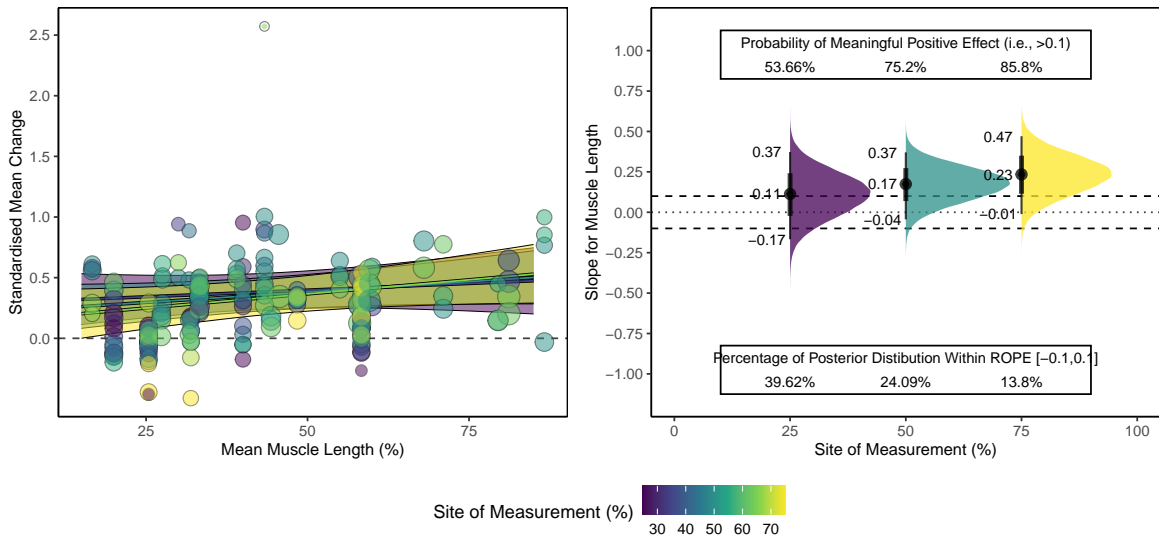
of 75.2% and percentage within the ROPE (i.e., -0.1,0.1) of 24.09%, and 0.23 [95% quantile interval: -0.01,0.47] at the 75% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 85.8% and percentage within the ROPE (i.e., -0.1,0.1) of 13.8%. There was considerable heterogeneity of effects relative to the magnitude of the population level effects, particularly at the study level, with $\tau_{study} = 0.09$ [95% quantile interval: 0.03,0.35], $\tau_{arm} = 0.05$ [95% quantile interval: 0.01,0.25], and $\tau_{effect} = 0.01$ [95% quantile interval: 0,0.07].

For the lnRR model the magnitude of muscle length slope was 1.57% [95% quantile interval: -4.28%, 7.89%] at the 25% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 32.66% and percentage within the ROPE (i.e., -0.1,0.1) of 60.86%, 3.18% [95% quantile interval: -1.67%,8.08%] at the 50% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 52.92% and percentage within the ROPE (i.e., -0.1,0.1) of 46.53%, and 4.81% [95% quantile interval: -0.84%,10.74%] at the 75% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 73.08% and percentage within the ROPE (i.e., -0.1,0.1) of 26.62%. There was considerable heterogeneity of effects relative to the magnitude of the population level effects, particularly at the study level, with $\tau_{study} = 4.21\%$ [95% quantile interval: 3.92%,14.59%], $\tau_{arm} = 1.98\%$ [95% quantile interval: 1.26%,7.47%], and $\tau_{effect} = 0.57\%$ [95% quantile interval: 0.36%,1.61%].

Model diagnostics can be seen in the supplementary materials here: <https://osf.io/up4q5>.

Interaction between mean muscle length and site of measurement

Global grand mean and 95% quantile intervals presented for predictions and slopes at 25%, 50%, and 75% of centred site of measurement

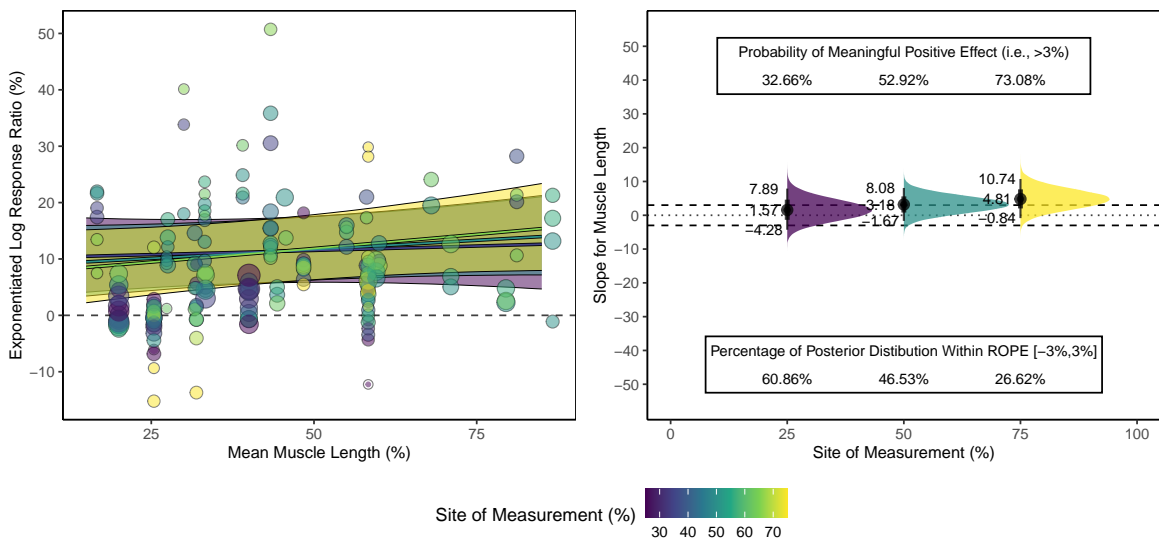


Note, the slopes have been transformed to the effect when increasing muscle length by 50% to reflect typical difference between short vs long lengths

Figure 3: Results from main model with uninformative priors for standardised mean difference effects.

Interaction between mean muscle length and site of measurement

Global grand mean and 95% quantile intervals presented for predictions and slopes at 25%, 50%, and 75% of centred site of measurement



Note, the slopes have been transformed to the effect when increasing muscle length by 50% to reflect typical difference between short vs long lengths

Figure 4: Results from main model with uninformative priors for exponentiated log response ratio effects.

Main Model - Addition of Random Slopes

For the main model utilising priors from Wolf et al. (2023), yet with the *addition* of random slopes for muscle length and measurement site at the study level, 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 50% e.g., the slope of the difference between 25% and 75% muscle length) at three levels of site of measurement (25%, 50%, and 75%) can be seen in Figure 5 for the SMD model, and Figure 6 for the lnRR model.

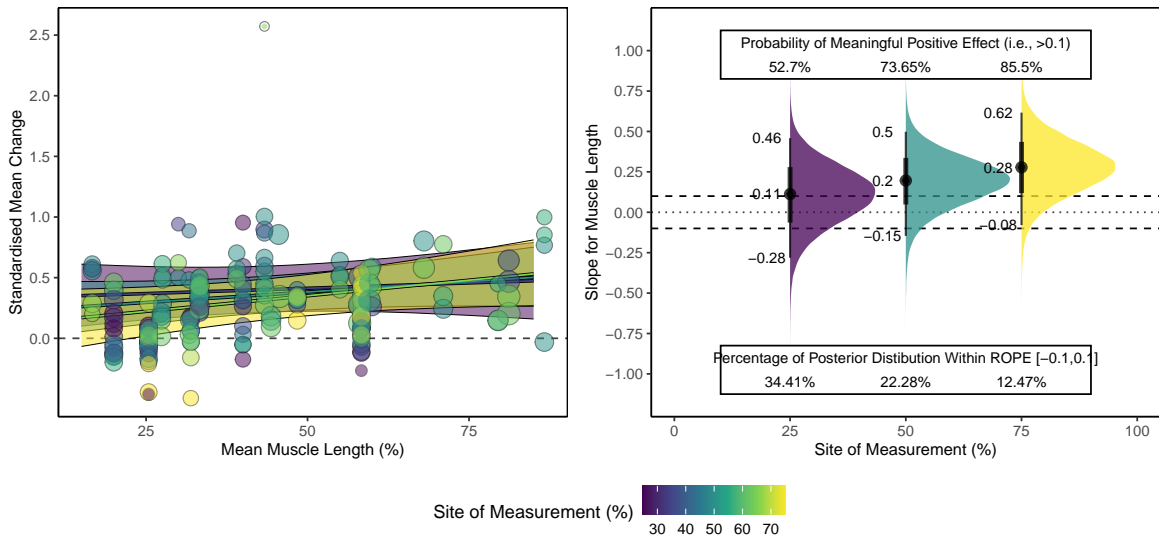
For the SMD model the magnitude of muscle length slope was 0.11 [95% quantile interval: -0.28, 0.46] at the 25% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 52.7% and percentage within the ROPE (i.e., -0.1,0.1) of 34.41%, 0.2 [95% quantile interval: -0.15,0.5] at the 50% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 73.65% and percentage within the ROPE (i.e., -0.1,0.1) of 22.28%, and 0.28 [95% quantile interval: -0.08,0.62] at the 75% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 85.5% and percentage within the ROPE (i.e., -0.1,0.1) of 12.47%. There was considerable heterogeneity of effects relative to the magnitude of the population level effects, particularly at the study level, with $\tau_{study} = 0.1$ [95% quantile interval: 0.03,0.4], $\tau_{arm} = 0.06$ [95% quantile interval: 0.01,0.26], and $\tau_{effect} = 0.01$ [95% quantile interval: 0,0.07]. At the study level there was considerable variation in the slopes too for muscle length $\tau_{study,muscle\ length} = 0.15$ [95% quantile interval: 0.01,0.46] and measurement site $\tau_{study,measurement\ site} = 0.18$ [95% quantile interval: 0.01,0.48]

For the lnRR model the magnitude of muscle length slope was 2.27% [95% quantile interval: -7.04%, 12.2%] at the 25% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 43.61% and percentage within the ROPE (i.e., -0.1,0.1) of 44.25%, 3.89% [95% quantile interval: -4.91%,13.15%] at the 50% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 58.77% and percentage within the ROPE (i.e., -0.1,0.1) of 35.9%, and 5.58% [95% quantile interval: -3.71%,15.86%] at the 75% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 72.2% and percentage within the ROPE (i.e., -0.1,0.1) of 24.47%. There was considerable heterogeneity of effects relative to the magnitude of the population level effects, particularly at the study level, with $\tau_{study} = 4.55\%$ [95% quantile interval: 3.49%,16.18%], $\tau_{arm} = 2.02\%$ [95% quantile interval: 0.8%,8.03%], and $\tau_{effect} = 0.68\%$ [95% quantile interval: 0.46%,1.74%]. At the study level there was considerable variation in the slopes too for muscle length $\tau_{study,muscle\ length} = 4.52\%$ [95% quantile interval: 0.59%,13.44%] and measurement site $\tau_{study,measurement\ site} = 3.22\%$ [95% quantile interval: 0.55%,8.82%]

Model diagnostics can be seen in the supplementary materials here: <https://osf.io/2b9dx>.

Interaction between mean muscle length and site of measurement

Global grand mean and 95% quantile intervals presented for predictions and slopes at 25%, 50%, and 75% of centred site of measurement

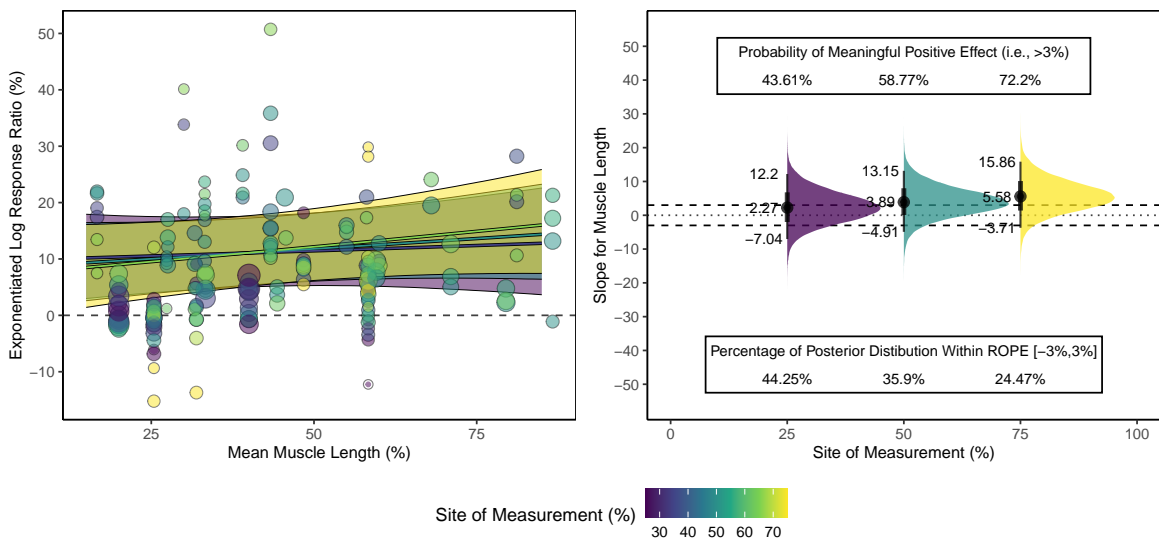


Note, the slopes have been transformed to the effect when increasing muscle length by 50% to reflect typical difference between short vs long lengths

Figure 5: Results from main model with addition of random slopes for standardised mean difference effects.

Interaction between mean muscle length and site of measurement

Global grand mean and 95% quantile intervals presented for predictions and slopes at 25%, 50%, and 75% of centred site of measurement



Note, the slopes have been transformed to the effect when increasing muscle length by 50% to reflect typical difference between short vs long lengths

Figure 6: Results from main model with addition of random slopes for exponentiated log response ratio effects.

Main Model with Random Slopes - J. Steele Priors

For the main model utilising the priors elicited from JS 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 50% e.g., the slope of the difference between 25% and 75% muscle length) at three levels of site of measurement (25%, 50%, and 75%) can be seen in Figure 7 for the SMD model, and Figure 8 for the lnRR model.

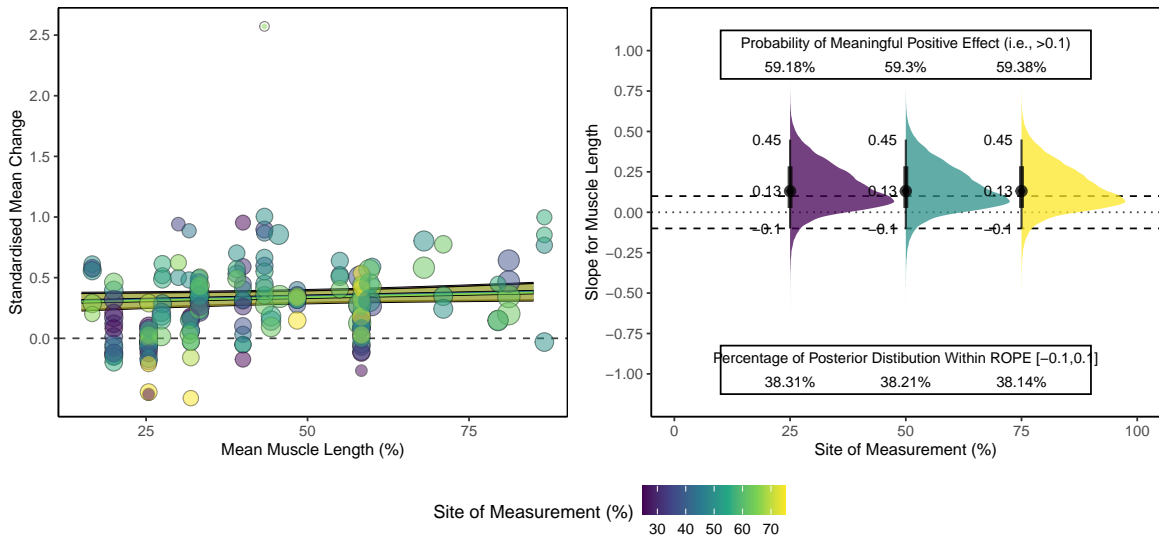
For the SMD model the magnitude of muscle length slope was 0.13 [95% quantile interval: -0.1, 0.45] at the 25% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 59.18% and percentage within the ROPE (i.e., -0.1,0.1) of 38.31%, 0.13 [95% quantile interval: -0.1,0.45] at the 50% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 59.3% and percentage within the ROPE (i.e., -0.1,0.1) of 38.21%, and 0.13 [95% quantile interval: -0.1,0.45] at the 75% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 59.38% and percentage within the ROPE (i.e., -0.1,0.1) of 38.14%. There was considerable heterogeneity of effects relative to the magnitude of the population level effects, particularly at the study level, with $\tau_{study} = 0.09$ [95% quantile interval: 0.1,0.27], $\tau_{arm} = 0.04$ [95% quantile interval: 0,0.16], and $\tau_{effect} = 0.01$ [95% quantile interval: 0,0.06]. At the study level there was considerable variation in the slopes too for muscle length $\tau_{study,muscle\ length} = 0.17$ [95% quantile interval: 0.01,0.47] and measurement site $\tau_{study,measurement\ site} = 0.15$ [95% quantile interval: 0.01,0.41]

For the lnRR model the magnitude of muscle length slope was 1.83% [95% quantile interval: -5.88%, 12.86%] at the 25% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 39.15% and percentage within the ROPE (i.e., -0.1,0.1) of 53.88%, 1.83% [95% quantile interval: -5.88%,12.86%] at the 50% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 39.17% and percentage within the ROPE (i.e., -0.1,0.1) of 53.87%, and 1.84% [95% quantile interval: -5.88%,12.86%] at the 75% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 39.21% and percentage within the ROPE (i.e., -0.1,0.1) of 53.84%. There was considerable heterogeneity of effects relative to the magnitude of the population level effects, particularly at the study level, with $\tau_{study} = 2.97\%$ [95% quantile interval: 2.12%,11.4%], $\tau_{arm} = 1.91\%$ [95% quantile interval: 0.38%,8.62%], and $\tau_{effect} = 0.42\%$ [95% quantile interval: 0.27%,1.19%]. At the study level there was considerable variation in the slopes too for muscle length $\tau_{study,muscle\ length} = 5.45\%$ [95% quantile interval: 0.45%,15.36%] and measurement site $\tau_{study,measurement\ site} = 2.53\%$ [95% quantile interval: 0.37%,6.87%]

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

Interaction between mean muscle length and site of measurement

Global grand mean and 95% quantile intervals presented for predictions and slopes at 25%, 50%, and 75% of centred site of measurement

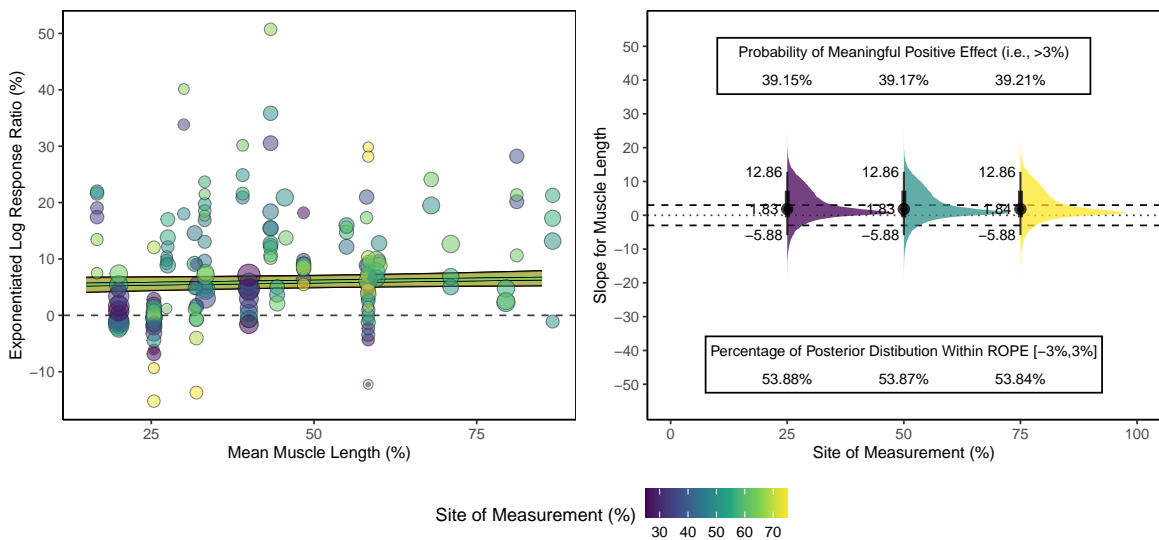


Note, the slopes have been transformed to the effect when increasing muscle length by 50% to reflect typical difference between short vs long lengths

Figure 7: Results from main model with J. Steele priors for standardised mean difference effects.

Interaction between mean muscle length and site of measurement

Global grand mean and 95% quantile intervals presented for predictions and slopes at 25%, 50%, and 75% of centred site of measurement



Note, the slopes have been transformed to the effect when increasing muscle length by 50% to reflect typical difference between short vs long lengths

Figure 8: Results from main model with J. Steele priors for exponentiated log response ratio effects.

Main Model with Random Slopes - Other Authors Priors

For the main model utilising the priors elicited from the other authors 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 50% e.g., the slope of the difference between 25% and 75% muscle length) at three levels of site of measurement (25%, 50%, and 75%) can be seen in Figure 9 for the SMD model, and Figure 10 for the lnRR model.

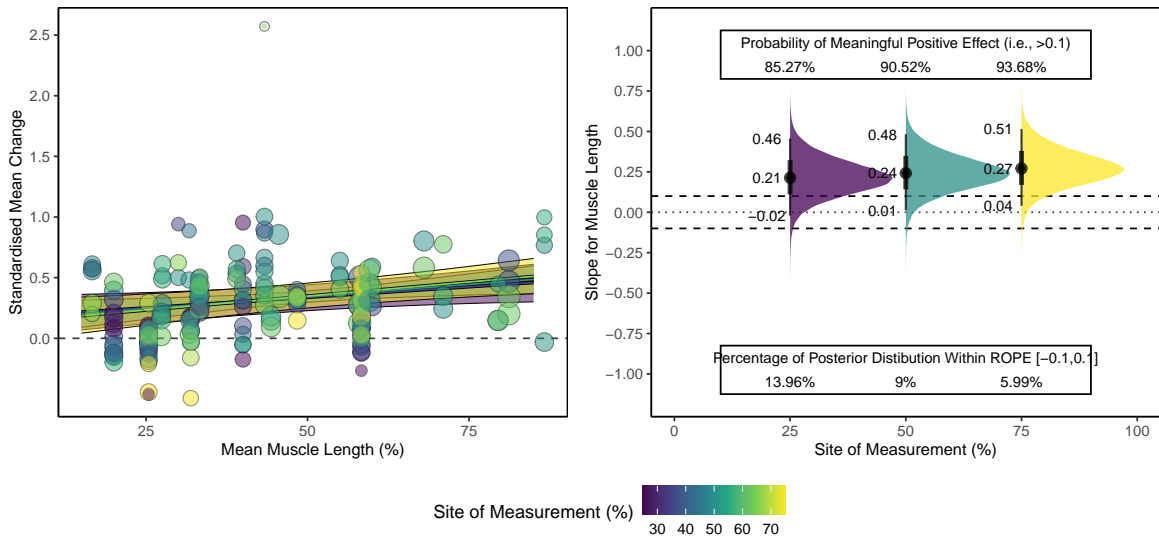
For the SMD model the magnitude of muscle length slope was 0.21 [95% quantile interval: -0.02, 0.46] at the 25% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 85.27% and percentage within the ROPE (i.e., -0.1,0.1) of 13.96%, 0.24 [95% quantile interval: 0.01,0.48] at the 50% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 90.52% and percentage within the ROPE (i.e., -0.1,0.1) of 9%, and 0.27 [95% quantile interval: 0.04,0.51] at the 75% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 93.68% and percentage within the ROPE (i.e., -0.1,0.1) of 5.99%. There was considerable heterogeneity of effects relative to the magnitude of the population level effects, particularly at the study level, with $\tau_{study} = 0.1$ [95% quantile interval: 0.12,0.28], $\tau_{arm} = 0.03$ [95% quantile interval: 0,0.14], and $\tau_{effect} = 0.01$ [95% quantile interval: 0,0.07]. At the study level there was considerable variation in the slopes too for muscle length $\tau_{study,muscle\ length} = 0.14$ [95% quantile interval: 0.01,0.42] and measurement site $\tau_{study,measurement\ site} = 0.19$ [95% quantile interval: 0.01,0.51]

For the lnRR model the magnitude of muscle length slope was 5.92% [95% quantile interval: -1.75%, 13.25%] at the 25% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 82.01% and percentage within the ROPE (i.e., -0.1,0.1) of 16.37%, 6.66% [95% quantile interval: -1.07%,14.06%] at the 50% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 87.19% and percentage within the ROPE (i.e., -0.1,0.1) of 11.52%, and 7.43% [95% quantile interval: -0.4%,15%] at the 75% measurement site with probability of a meaningful positive effect (i.e., 0.1) of 90.54% and percentage within the ROPE (i.e., -0.1,0.1) of 8.41%. There was considerable heterogeneity of effects relative to the magnitude of the population level effects, particularly at the study level, with $\tau_{study} = 3.81\%$ [95% quantile interval: 2.75%,12.49%], $\tau_{arm} = 1.05\%$ [95% quantile interval: 0.35%,6.69%], and $\tau_{effect} = 0.47\%$ [95% quantile interval: 0.32%,1.25%]. At the study level there was considerable variation in the slopes too for muscle length $\tau_{study,muscle\ length} = 4.79\%$ [95% quantile interval: 0.52%,13.2%] and measurement site $\tau_{study,measurement\ site} = 3.83\%$ [95% quantile interval: 0.53%,9.69%]

Model diagnostics can be seen in the supplementary materials here: <https://osf.io/g2e57>.

Interaction between mean muscle length and site of measurement

Global grand mean and 95% quantile intervals presented for predictions and slopes at 25%, 50%, and 75% of centred site of measurement

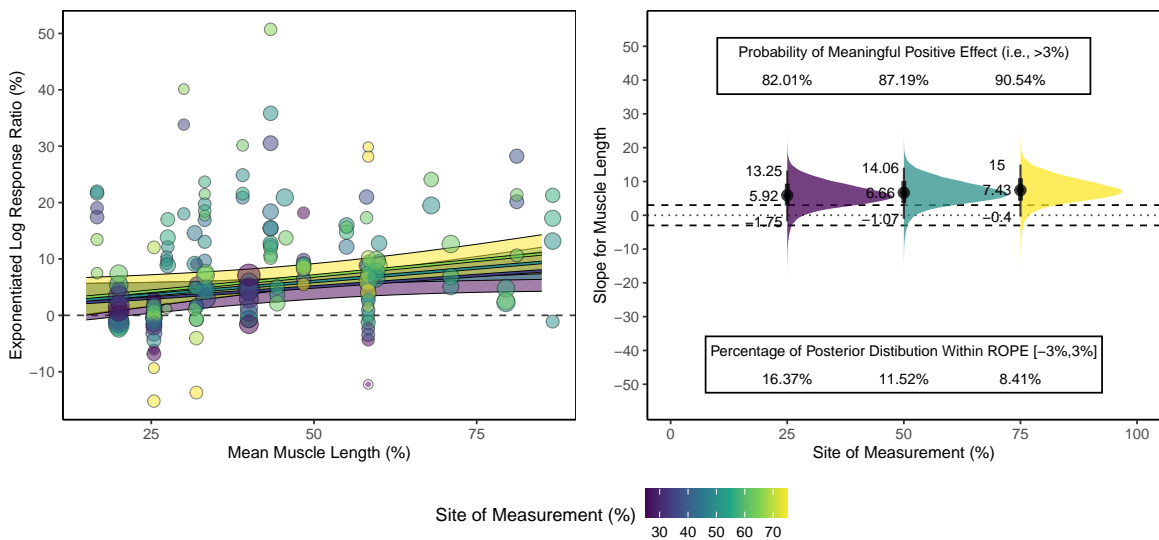


Note, the slopes have been transformed to the effect when increasing muscle length by 50% to reflect typical difference between short vs long lengths

Figure 9: Results from main model with other authors priors for standardised mean difference effects.

Interaction between mean muscle length and site of measurement

Global grand mean and 95% quantile intervals presented for predictions and slopes at 25%, 50%, and 75% of centred site of measurement



Note, the slopes have been transformed to the effect when increasing muscle length by 50% to reflect typical difference between short vs long lengths

Figure 10: Results from main model with other authors prior for exponentiated log response ratio effects.

Model Comparisons

The comparison between all the models presented for the SMD models can be seen in Figure 11, and for the lnRR models in Figure 12. The uninformative SMD model had only weak evidence against though the lnRR positive evidence against when compared to the pre-registered model utilising priors from Wolf et al. (2023). Interestingly the addition of random slopes to the main model fared worse than when they were fixed between units within the study level for both the SMD model with strong evidence against, and the lnRR model with very strong evidence against. Both of the two models using more informative priors fared much better than all other models with strong to very strong evidence to support them for both SMD and lnRR models; though the model utilising the priors selected by JS had weak evidence favouring it over the model utilising priors selected by the other authors for both SMD and lnRR models.

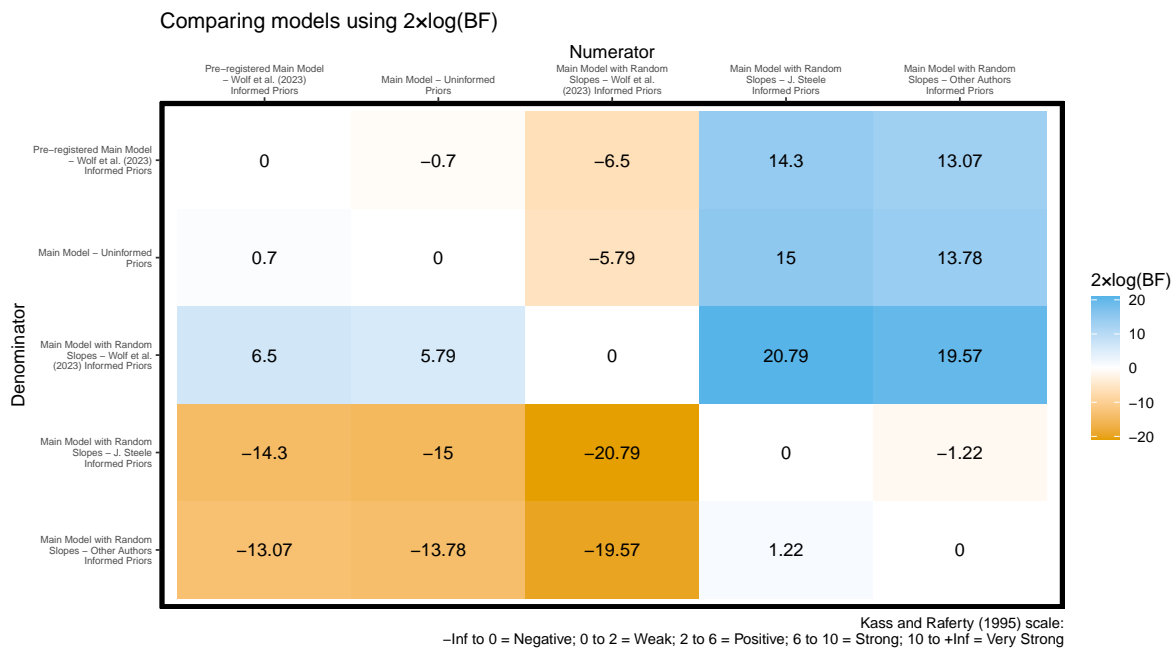


Figure 11: Model comparison using Bayes Factors for SMD models.

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>
- Dragicevic, P., Jansen, Y., Sarma, A., Kay, M., & Chevalier, F. (2019). Increasing the Transparency of Research Papers with Explorable Multiverse Analyses. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. <https://doi.org/10.1145/3290605.3300295>

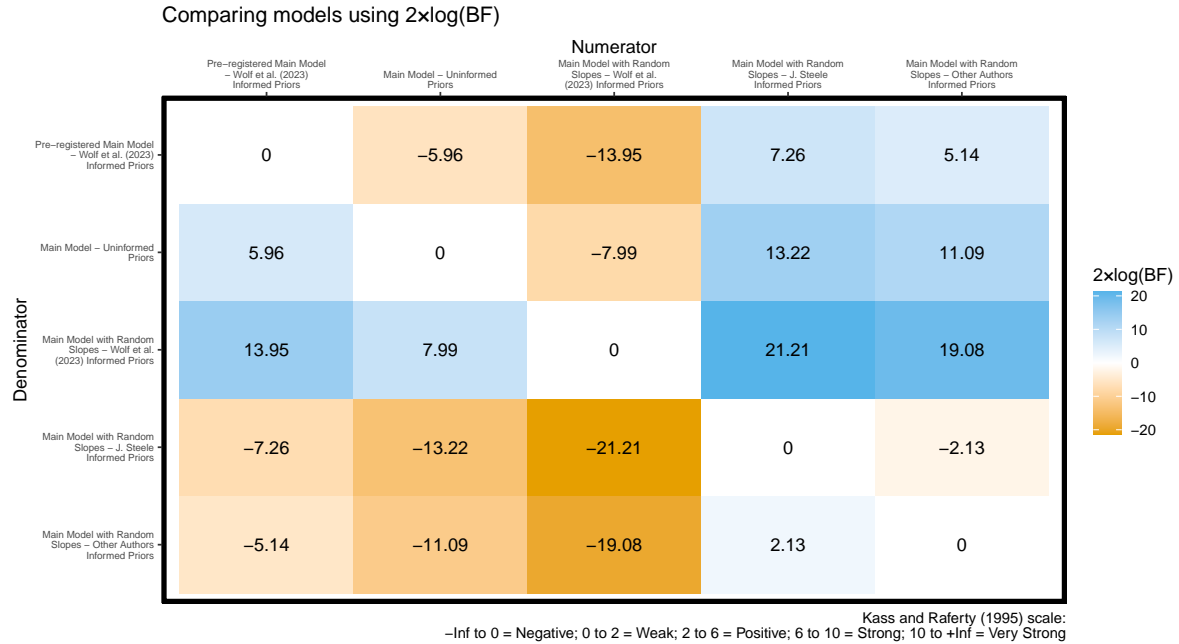


Figure 12: Model comparison using Bayes Factors for lnRR models.

- 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>
- Kallioinen, N., Paananen, T., Bürkner, P.-C., & Vehtari, A. (2023). Detecting and diagnosing prior and likelihood sensitivity with power-scaling. *Statistics and Computing*, 34(1), 57. <https://doi.org/10.1007/s11222-023-10366-5>
- Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, 90(430), 773–795. <https://doi.org/10.1080/01621459.1995.10476572>
- 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>
- 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>

[org/10.1098/rsif.2017.0213](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>
- Zabaleta-Korta, A., Fernández-Peña, E., & Santos-Concejero, J. (2020). Regional Hypertrophy, the Inhomogeneous Muscle Growth: A Systematic Review. *Strength & Conditioning Journal*, 42(5), 94. <https://doi.org/10.1519/SSC.0000000000000574>