Is There Too Much of a Good Thing? Meta-Regressions of the Effect of Per-Session Volume on Hypertrophy and Strength

Supplementary materials:
https://osf.io/dqka3/
For correspondence:
jremmert2020@fau.edu

Jacob F. Remmert[1], Joshua C. Pelland[1], Zac P. Robinson[1], Seth R. Hinson[1], Michael C. Zourdos[1]

[1] Department of Exercise Science and Health Promotion, Muscle Physiology Laboratory, Florida Atlantic University

Please cite as: Remmert, J., Pelland, J., Robinson, Z., Hinson, S., Zourdos, M. (2025). Is There Too Much of a Good Thing? Meta-Regressions of the Effect of Per-Session Volume on Hypertrophy and Strength. *SportRχiv.*

All authors have read and approved this version of the manuscript. This article was last modified in March 2025.

ABSTRACT

Background: Recent research has quantified the dose-response relationship between weekly resistance training set volume and muscle hypertrophy and strength gains. However, the nature of the dose-response of *per-session* set volume remains underexplored.
Objective: Before meta-analyzing, all contributing sets were classified as direct or indirect, depending on their specificity to the measurement. Then, per-session set volume for indirect sets was quantified as 1 for 'total,' 0.5 for 'fractional,' and 0 for 'direct.' A series of multi-level meta-regressions were performed for muscle hypertrophy and strength, with all models adjusted for intervention duration and training status. The point of undetectable outcome superiority (PUOS) was identified as the per-session set volume in which additional sets did not result in a >50% likelihood of the difference in hypertrophy or strength gain exceeding the smallest detectable effect size.
Results: The 'direct' and 'fractional' quantification methods provided the strongest relative evidence for strength and hypertrophy, respectively; thus, these quantification methods were used for the primary meta-regression models. The posterior probability of the marginal slope exceeding zero for the effect of per-session volume on both strength and hypertrophy was 100%, indicating positive dose-response relationships between per-session set volume and hypertrophy and strength gains. However, both best fit models suggest diminishing returns as per-session set volume increases, with the PUOS occurring at ~2 'direct' sets for strength and ~11 'fractional' sets for hypertrophy. Notably, the 'direct' set model for strength gains suggests more strongly diminishing returns compared to the 'fractional' set model for hypertrophy.
Conclusions: There is a positive dose-response relationship between per-session volume with both strength and hypertrophy; however, to quantify the dose-response relationship, it is paramount to distinguish between 'fractional' and 'direct' set counting methods. Furthermore, the relationship exhibits diminishing returns for both outcomes, which appear to manifest more strongly for strength gains compared to hypertrophy. While the available evidence indicates a PUOS of ~2 and ~11 sets per session for strength gain and hypertrophy, respectively, there is insufficient data with very high per-session set volumes. Therefore, it is unclear whether there is a point in which additional per-session sets attenuate adaptations, or if even higher per-session set volumes could be potentially beneficial; thus, the PUOS should be interpreted cautiously.

1 INTRODUCTION

When designing a resistance training (RT) program, organizing volume (number of sets) and frequency (the number of times per week a muscle or exercise is trained) appropriately may impact training-induced adaptations. Various meta-analyses have explored the effects of weekly volume (1–5) and frequency (6–11) on muscle hypertrophy and strength gain. The most recent meta-analytic data, a parallel project from our laboratory (2), found a meaningful and positive dose-response relationship between weekly set volume and both muscle hypertrophy and strength gain. However, there were diminishing returns, with no detectable superiority beyond ~31 'fractional' sets per week for hypertrophy and ~3 'fractional' sets per week for strength gain. Pelland et al. (2) also reported a positive dose-response between frequency and strength gain. At the average frequency in the dataset, the marginal slope indicated a 3.27% (95% CrI: 2.74%, 3.84%) increase in strength gain with each additional training day. Conversely, the marginal slope of the dose-response between frequency and hypertrophy was only 0.32% (95% CrI: -0.14%, 0.82%). While this effect was positive, it was small and fell within the typical range of error seen in RT studies (12).

Integrating the independent effects of weekly set volume and frequency from Pelland et al. (2) raises important questions. For hypertrophy, the relatively strong effect of weekly set volume combined with the unclear effect of frequency leaves the optimal per-session set volume unspecified. While there was undetectable outcome superiority beyond ~31 'fractional' weekly sets, the absence of a consistent frequency effect—beyond typical errors in RT studies—further complicates practical recommendations. Since performing 31 sets in a single session is impractical, identifying appropriate per-session volumes remains crucial. For strength gains, Pelland et al. (2) reported undetectable superiority beyond ~3 fractional weekly sets and beyond a 'fractional' frequency of 3. This suggests that low per-session set volumes may be sufficient to maximize strength gains; however, a direct analysis is needed for confirmation.

Although less common than analyses of weekly set volume, some research investigating per-session set volume exists. A 2010 meta-analysis by Krieger (3) reported greater hypertrophy with multiple sets per exercise compared to a single set (ES = 0.10 [0.02, 0.19]; p = 0.016). More specific comparisons (2-3 sets per exercise vs. 4-6 sets per exercise) tended to indicate greater effects with higher set volumes, but none reached statistical significance. However, quantifying volume as sets per exercise may lack the specificity to accurately assess the dose-response relationship of per-session set volume. In contrast to the findings from Krieger (3), a 2020 meta-analysis by Benito et al. (13) had a negative effect

on combined assessments of lean mass (fat free mass, lean muscle mass, and skeletal muscle mass) (β = -0.03 [-0.05–0.001], p = 0.04), potentially suggesting an upper limit of per-session volume beyond which changes in lean mass are negative. However, Benito et al. utilized a wide inclusion criteria and didn't necessarily analyze studies with *differences* in per-session volumes. Further, both Krieger (3) and Benito et al (13) included indirect measures of muscle size (e.g., dual-energy X-ray absorptiometry [DXA]), arm circumference), which may not be as sensitive as direct, site-specific measures (e.g., ultrasonography, MRI, and muscle biopsy) of hypertrophy.

In a later unpublished 2017 meta-regression (14), Krieger examined the effect of per-session set volume per muscle group as a continuous variable on hypertrophy using the same extracted data as Schoenfeld et al. (5) (i.e., 34 intervention groups from 15 studies). Krieger qualitatively determined an apparent point of diminishing returns for hypertrophy using only studies with interset rest times ≥2 minutes and reported an upper threshold of approximately 6-8 sets per-session. However, this qualitative analysis may have indicated the *most likely* breakpoint instead of necessarily providing strong evidence of a clear breakpoint. This, along with the lack of a clear upper limit when analyzing all studies in Krieger's analysis, suggests a wider range of functional forms should be explored. No previous per-session meta-analysis has analyzed only direct, site-specific hypertrophy measures, quantified only sets per session for the measured muscle group (not per exercise), and treated set volume as a continuous variable (15,16). Further, all previous meta-analyses have used a volume quantification method similar to the 'total' method reported by Pelland et al. (2), in which all sets for a muscle group are quantified the same, even if the measured muscle was unlikely to be the primary force generator. Pelland et al. (2) distinguished between direct sets (i.e., the measured muscle was likely to be the primary force generator) and indirect sets (i.e., the measured muscle likely contributed meaningfully to force production but was unlikely to be the primary force generator). Indeed, the relative evidence for the 'fractional' quantification method – in which indirect sets were counted as half a set - had stronger relative evidence for hypertrophy (Bayes Factor = 9.48) than the 'total' quantification method.

To our knowledge, no data are currently available investigating the effects of per-session set volume on strength gain. Based upon the results of Pelland et al. (2), one might reasonably hypothesize very strong diminishing returns with higher per-session set volume, with very few sets in each of 2-3 sessions per week comprising an appropriate RT program. However, a direct analysis on the potential dose-response effect of sets per session on strength gain is needed to gain clarity on the topic.

Therefore, the purpose of this analysis was to utilize meta-regression to investigate the effects of per-session volume while quantifying 'indirect' sets in various ways as a continuous variable on direct, site-specific measures of hypertrophy and measures of maximal strength. We hypothesized a meaningful and positive dose-response relationship with greater volumes leading to greater adaptations, although with diminishing returns. We also examined a series of moderator analyses in an exploratory manner without formal hypotheses.

2 METHODS

This meta-analysis was conducted according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (17). Pre-registration on the Open Science Framework (https://osf.io/dqka3/) using the International Prospective Register of Systematic Reviews template was used, though some of the methods have changed since the original pre-registration.

2.1 Quality of Evidence Assessment

We used the TESTEX scale (18) to assess the quality of included studies due to its design specifically for exercise training studies. To aid in interpretation, we also included 1) funnel plots and effective sample size approximated bias-adjusted estimates (19) for assessing heterogeneity and small study bias (20), 2) contrast-based meta-analyses to assess the consistency of effects between higher vs. lower per-session volume, 3) additional meta-regressions with strict inclusion criteria, described in section 3.

2.2 Inclusion and Exclusion Criteria

English-language full-text peer reviewed articles as well as pre-prints and masters and doctoral theses were considered for inclusion. Inclusion criteria consisted of a randomized experimental design with the presence of a dynamic resistance training intervention lasting at least four weeks and comparing at least two conditions or groups featuring differences in training volume (defined as the number of sets per week) or frequency (defined as the number of sessions per week) or both. Further, to be included, studies must have controlled for load (±5% of 1 repetition maximum (RM) or ±2RM), exercise selection, and proximity to failure (failure or nonfailure or mixed) and reported outcome measures at pre-testing and post-testing for maximal strength (isometric, isokinetic, or isotonic strength up

to a 10RM) or hypertrophy (direct measures only: ultrasound, computed tomography (CT), magnetic resonance imaging (MRI), muscle biopsy). No retracted studies or studies called into question by Vigotsky et al. (21) were included.

Exclusion criteria included studies involving diseased or unhealthy subjects or subjects greater than 70 yrs old (if not explicitly stated, if mean age plus two times the standard deviation was greater than 70, that was considered grounds for exclusion). Additionally, studies involving a supplement intervention were excluded.

2.3 Literature Search and Screening

PubMed and Google Scholar databases were searched for English-language studies for all time points up until April 2023. Studies published between April 2023 and June 2024 that the authors became aware of were also considered for inclusion. The literature search terms, screening process, and extraction process are described in our laboratory's parallel project (2).

2.4 Data Extraction and Coding

Data relevant to strength and hypertrophy outcomes were extracted from the included studies, including details regarding sample characteristics, training interventions, and the outcome(s). When data were not available, the authors were contacted to request the missing data. If no response was received, a second request was sent. If there was still no response, data were obtained using WebPlotDigitizer (v5.0, Ankit Rohatgi) where possible. Descriptive data were extracted as separate variables for total, direct, and fractional volume quantification methods. Data extraction and coding were performed by JR, JP, and SH.

2.5 Volume Quantification

Per-session set volume was quantified in three ways: 'direct,' 'total,' and 'fractional' sets.

For hypertrophy outcomes, 'direct' sets included only sets of exercises in which the likely primary force generator is the measured muscle; for example, if the biceps brachii are measured, only sets of isolated elbow flexion were counted for 'direct' sets. 'Total' sets included all 'direct' sets plus sets where the measured muscle likely contributes as a synergist; for example, if the biceps brachii are measured, sets of isolated elbow flexion

and sets of non-isolated elbow flexion, such as lat pulldowns, are counted equally. 'Fractional' sets included the same sets as 'total' sets but sets where the measured muscle was not the likely primary force generator were counted as half of a set; for example, if the biceps brachii are measured, sets of isolated elbow flexion count as one set while sets of non-isolated elbow flexion count as half a set.

For strength outcomes, 'direct' sets included only sets performed of the same exercise as the tested strength assessment; for example, if barbell back squat is assessed, only barbell back squat sets are counted. 'Total' sets included all 'direct' sets plus any sets that were likely to meaningfully train the muscle(s) involved in the strength assessment; for example, if barbell back squat is assessed, barbell back squat sets and leg press sets are counted equally. 'Fractional' sets included the same sets as 'total' sets, but indirect sets were counted as half of a set; for example, if barbell back squat is assessed, barbell back squat sets are counted as one set and leg press sets are counted as half a set.

This process was not wholly objective; therefore, a tabulated breakdown of direct versus indirect set classification can be found in Table 1 of our parallel project (2).

2.6 Control Group Estimates

To appropriately contextualize model estimates and inform the smallest detectable effect size (SDES), we sought to include data from non-training control groups of untrained participants from the included studies. However, as only 13 included studies (22–34) featured such data, with only two of these (23,34) having hypertrophy outcomes, this became inadvisable. Thus, we instead utilized published data from Steele et al. (12) which meta-analyzed 368 strength effects and 223 hypertrophy effects to inform the SDES (specifically, the square root of the sum of the estimated variance components from the multi-level meta-analytic model), resulting in values of 2.05% and 3.96% for hypertrophy and strength outcomes, respectively (https://osf.io/qzxcj).

3 Statistical Analyses:

This meta-analysis was performed using the *brms* and *metafor* packages in the R language and environment for statistical computing (v 4.0.2; R Core Team, https://www.r-project.org). Available on the Open Science Framework (https://osf.io/dqka3/) are the extracted data, analysis scripts, estimates, plots, and supplementary files. Due to the nature of this project, we did not conduct null hypothesis significance testing to avoid dichotomizing our findings (35), and instead used an approach based on estimation within a Bayesian framework in

which effect estimates and their precision were interpreted continuously and probabilistically (36). Nested effect sizes were calculated due to many of the included studies consisting of multiple effects reported for each of multiple groups.

The primary analyses were performed as multilevel arm-based meta-regression models (37,38) with study, group, and observations as nested random intercepts in the model (observations nested within groups and groups within studies). We attempted to include random slopes on the study-level for the volume variable to account for heterogeneity in the fixed effects across studies, however in nearly all cases this resulted in unresolvable model warnings, such as divergent transitions. Therefore, random slopes were not included. Effects were weighted by inverse sampling variance to account for variance between studies, within studies, and between observations in a study.

3.1 Effect Size Calculation

Following model construction, effect sizes and associated variance were calculated using the *escalc* function (39,40) as both response ratios and standardized mean change, since the effect size calculation choice may influence model selection (41). Response ratios were calculated as the sum of the natural logarithm of the ratio of post-test and pre-test means, which were later exponentiated (i.e., e^x) to convert them to percentage change scores to make practical interpretation more intuitive. Standardized mean changes were calculated as the difference between post-test and pre-test means, divided by the pooled pre-test standard deviation with an adjustment (i.e., C) for small sample bias. Formulas for each effect size and their variances can be seen below:

$$SMC = C\left(\frac{Mean_{post} - Mean_{pre}}{SD_{pre}}\right) \quad ; C = 1 - \left(\frac{3}{4(n-1) - 1}\right)$$

$$var(SMC) = \frac{2(1 - r)}{n} + \frac{(SMC)^2}{2 \cdot n}$$

$$RR = ln\left(\frac{Mean_{post}}{Mean_{pre}}\right)$$

$$var(RR) = \frac{(SD_{post})^2}{n \cdot (Mean_{post})^2} + \frac{(SD_{pre})^2}{n \cdot (Mean_{pre})^2} - \frac{2 \cdot r \cdot (SD_{post}) \cdot (SD_{pre})}{n \cdot (Mean_{post}) \cdot (Mean_{pre})}$$

$$RR_{exp} = (e^{lnRR} - 1) * 100$$

Regarding calculating the variance of effect sizes, most studies did not report the necessary pre-testing to post-testing correlation values. If possible, available data were used to estimate those values to be used for analysis (42); if not possible, values were imputed based on the median of the estimated correlation values. Similarly, if the standard deviations needed to calculate the effect sizes were missing, approximation methods were used via referencing a weighted coefficient of variation (43). To quantify the variance explained by only the fixed effects or the combined fixed and random effects, we calculated marginal and conditional R^2 values (44).

3.2 Best Fit Model Selection

Because of the potential of a nonlinear dose-response relationship between volume and strength or hypertrophy, we preliminarily fit a variety of functional forms for all models with the *metafor* package as random intercept models with both effect size types, and compared them using the *performance* package (45). Bayesian Information Criterion (BIC) and an approximated Bayes Factor were used to determine which model led to the highest probability of the observed data for each outcome. The Bayes Factor was calculated for each model relative to an intercept-only model and subsequently averaged between effect sizes. The model that performed the best after accounting for effect sizes on both additive and multiplicative scales was considered the "best fit;" however, although both effect size calculations were used for model selection, only response ratio models are reported. The following functional forms were compared to find the "best fit model" for each outcome:

1. Linear
2. Restricted Cubic Spline (4 knots)
3. Linear-log
4. 2nd Order Polynomial
5. Square Root
6. Quadratic Term
7. Reciprocal

All models included per-session volume ('total,' 'fractional,' or 'direct'), 2) intervention duration (continuous; in weeks), and 3) training status of the participants (binary categorical; trained or untrained).

3.3 Estimates

Using the *emmeans* package (46), marginal effects (means) with 95% compatibility intervals (quartile-based credible and prediction intervals) were extracted for the main effect of per-session volume, adjusted proportionally for all other predictors. To better estimate the magnitude of all model predicted effect sizes, each presented estimate has been control-adjusted, meaning the mean effect size predicted at zero sets per session was contrasted with the mean effect size predicted at every other number of sets. Therefore, the models represent effect sizes and compatibility intervals of a given per-session volume relative to the control effect.

3.4 Moderator Analyses

Interaction moderator analyses were then performed with the *metafor* package for the best fit models to examine the influence of a variety of training- or participant-related factors, such as sex, age, proximity to failure, rest interval length, and more. For these analyses, separate models were fit for each moderator with the same structure as the best fit model, but with an additional linear main effect and interaction term between per-session volume and the moderator of interest. Given differences in the number of effects between levels of the moderator, non-training control groups and effects from studies that manipulated both volume and frequency were not included to ensure undue weight was not provided to these effects.

3.5 Alternative Meta-Analytic Approaches

As stated previously, to address the research question of the effects of per-session volume on strength and hypertrophy, this analysis includes studies that manipulate training volume, training frequency, or both. By including more data in this way, statistical power/precision is improved (47), and these differences in study design are accounted for in the multilevel structure and fixed effects of the primary models.

However, we additionally conducted an analysis including only volume- or frequency-manipulated studies, fitting traditional multilevel contrast-based models using between-condition effect sizes. An intercept-only model comparing higher to lower per-session volume, as well as a two-stage fixed effect meta-regression model were each examined. Regarding the latter, the first stage involved fitting independent sample size weighted linear models for each outcome in each study, then extracting the intercept and slope for each and pooling them for each study. Next, the estimates were weighted by sample size and pooled for each study in a multivariate model, accounting for residual correlations

between intercepts and slopes. The second stage involved meta-analyzing the intercepts and slopes across studies again, accounting for residual correlation between estimates by using a multivariate model. Finally, dose-response predictions were created via posterior distributions of the pooled intercept and slope. Due to the limited number of observations per regression model in stage one, a linear form was used for both stages of this analysis.

These alternative approaches were used to confirm a lack of undue bias in the primary meta-regression models via similar directionality and magnitude in the alternative models. Notably, these alternative models contain far less data and therefore provide less precision, and also may lose the benefits of effect regularization from partial pooling and shrinkage featured in the primary models. There was not an irreconcilable contradiction in the results of any of the best fit models; however, these models potentially aid in interpretation.

3.6 Comparing Volume Quantification Methods

To evaluate the relative evidence for each per-session volume quantification method ('total,' 'direct', and 'fractional'), we calculated Bayes Factors for each pairwise comparison between primary meta-regression models, resulting in a measure of which model ('total,' 'direct', and 'fractional') was the most probable given the included data. The strength of evidence was interpreted based on the Kass and Raferty scale (48).

3.7 Point of Undetectable Outcome Superiority

To aid practical interpretation, we estimated a point of undetectable outcome superiority (PUOS) for the effect of per-session set volume on strength and hypertrophy for the most probable meta-regression model. This was defined as the value of per-session volume at which no pairwise comparisons to higher per-session volume values showed a >50% probability of the cumulative differences to exceed the smallest detectable effect size.

In an attempt to aid practical interpretation, we converted the identified value of the PUOS from the favored model type into the other volume quantification methods. In other words, this conversion provides an approximation of the per-session set volume where there is <50% confidence that that additional hypertrophy continues to occur. To make this approximation, we first calculated the mean of total sets per session and indirect sets per session for each outcome in each included study. From those values, the proportion of indirect sets was calculated, then this proportion was used to convert the PUOS value to

the other set quantification types. For example, the PUOS for hypertrophy was 11 'fractional' sets. Using the calculated proportion of indirect sets, approximately 25.776%, the PUOS in terms of 'direct' sets was estimated as (11 * (1 - 0.25776)), resulting in a value of 8.165 'direct' sets. The remaining PUOS values for each set model for both outcomes can be seen in Table 1.

4 RESULTS

4.1 Literature Search

As detailed in Figure 1 in our parallel project on weekly set volume (2), a total of 6677 studies were identified by the literature search, and 162 were identified as duplicates and deleted, leaving 6515 studies for consideration. An additional 16 studies were identified through other means (i.e., authors becoming aware of a new publication, citation searching in relevant studies). 135 publications remained for consideration following title and abstract screening, whereupon the full texts were assessed for inclusion, ultimately resulting in 67 studies included in the analysis.

4.2 Quality of Evidence Assessment

The mode TESTEX score was 12/15 (range = 8-14; https://osf.io/9wzb6). The mode study quality score was 3/5 (range = 1-5). The mode study reporting score was 8/10 (range = 4-10). Qualitative assessment of funnel plots and effective sample size approximated bias-adjusted estimates did not indicate small-study bias for hypertrophy (https://osf.io/7g6xr) or strength (https://osf.io/xmgk5). Contrast-based meta-analyses and two-stage meta-regressions returned no consistent indication of larger-than-expected heterogeneity in results.

4.3 Study Characteristics

A detailed breakdown (https://osf.io/y97wg) and visual summary (Figure 2 of our parallel project on weekly set volume (2)) of the included studies are available. Briefly, our analysis included data from 2058 participants of a mean age 25.16 ± 5.22 yrs and intervention lasting 10.42 ± 4.48 weeks. Mean training descriptive data consisted of 13.00 ± 8.87 sets per week, 2.33 ± 0.98 sessions per week, 1.80 ± 0.68 min rest periods, and 10.63 ± 3.53 repetitions per set for hypertrophy effects, and 8.14 ± 6.23 sets per week, 1.97 ± 0.92

sessions per week, 2.04 ± 0.79 min rest periods, and 9.85 ± 3.19 repetitions per set for strength effects.

4.4 Volume Quantification Methods

For 'total,' 'direct,' and 'fractional' volume – for both hypertrophy and strength – separate meta-regression models were created, resulting in six total models. The relative strength of evidence for each model regarding hypertrophy and strength separately were assessed in terms of 2 × Log(BF) values (48) (https://osf.io/f7659). Regarding the hypertrophy models, there was 'positive' evidence that 'fractional' outperformed 'total' (Log(BF)=3.72) and 'weak' evidence that 'fractional' outperformed 'direct' (Log(BF)=0.96). Regarding the strength models, there was 'strong' evidence that 'direct' outperformed both 'fractional,' (Log(BF)=9.87) and 'total' (Log(BF)=6.95).

Given the most favorable evidence for the 'fractional' sets model for hypertrophy outcomes and the 'direct' sets model for strength outcomes, the following results sections will focus on those model results. Results for other models can be found in the supplementary materials (https://osf.io/dqka3/).

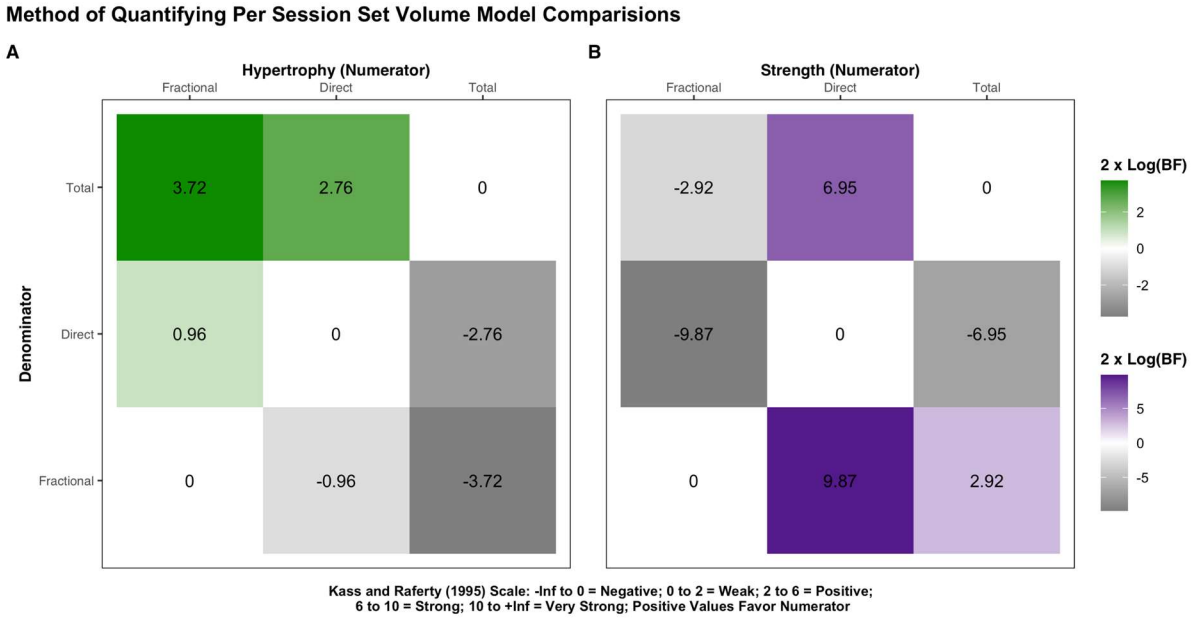

Figure 1. *Relative evidence for models using the Kass and Raftery scale: < 0 = negative evidence in favor of the numerator; 0 < 2 = weak evidence in favor of the numerator; 2 < 6 = positive evidence in favor of the numerator; 6 < 10 = strong evidence in favor of the numerator; ≥ 10 = very strong evidence in favor of the numerator.*

4.5 Primary Analysis

In the following sections, primary analysis results are presented regarding the meta-regression models for the effects of per-session volume on hypertrophy and strength. In particular, the best fit of all candidate models is presented alongside the overall quality of model fit (R^2) and the marginal slope for the main effect of per-session volume (i.e., the slope at the mean of per-session volume after adjusting for intervention duration and training status). Full model summary tables with all extracted estimates can be found in the supplementary materials (https://osf.io/dqka3/).

4.5.1 Hypertrophy Outcomes

Fractional Per-Session Set Volume:

The multilevel meta-regression models for the effect of 'fractional' per-session volume on hypertrophy outcomes included 220 total effects from 35 studies involving 1032 participants. Model comparisons revealed the linear-log model was the best fit (Figure 2). The fixed effects of the model explained approximately 16% of the variance ($R^2_{marginal}$ = 16.1%, $R^2_{conditional}$ = 73.8%). The marginal slope was positive with a 100% probability the mean linear slope exceeded zero, and the 95% credible interval did not contain a null point estimate (β = 0.393% [95% CrI: 0.202%, 0.583%]). The best fit model and slope indicate a positive dose-response relationship between hypertrophy and 'fractional' per-session volume.

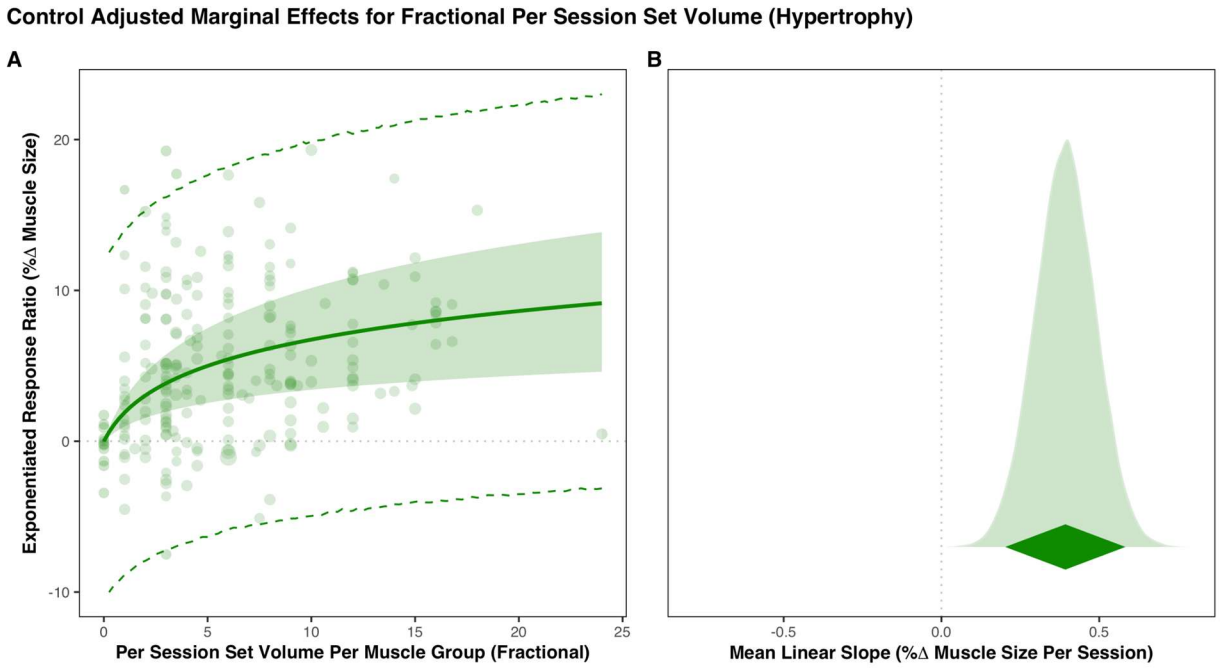

Figure 2: *Fractional per-session set volume best fit multilevel meta-regression for hypertrophy (linear-log model) analyzed as an exponentiated response ratio. Data are presented as estimated marginal means (solid line) with 95% quantile based compatibility intervals (light band = credible, dotted band = prediction) after adjusting for intervention duration and training status. Colored circles represent the effect size of each observation included in the analysis, with the size of each circle representing its weight determined by inverse variance weighting. Panel B represents the linear slope at the mean value of fractional per-session volume for all effects. In all panels, the main effect for fractional per-session volume is presented at the mean of the continuous fixed effect (i.e., intervention duration) and proportionally marginalized across the categorical fixed effect (i.e., training status).*

Additional models containing only direct effects (i.e., the intercept-only model comparing dichotomized higher versus lower per-session volumes and the meta-regression model of absolute per-session volume differences), including 121 total effects from 17 studies involving 544 participants, can be found in the supplementary materials (https://osf.io/6s9ym). These models qualitatively confirm the main meta-regression model.

4.5.2 Strength Outcomes

Direct Per-Session Set Volume:

The multilevel meta-regression models for the effect of direct per-session volume on strength outcomes included 490 total effects from 66 studies involving 2020 participants. Model comparisons revealed the reciprocal model was the best fit (Figure 3). The fixed effects of the model explained approximately 15% of the variance ($R^2_{marginal}$ = 14.9%, $R^2_{conditional}$ = 73.3%). The marginal slope was positive with a 100% probability the mean linear slope exceeded zero, and the 95% credible interval did not contain a null point estimate (β = 1.23% [95% CrI: 0.941%, 1.52%]). The best fit model and slope indicate a positive dose-response relationship between strength and 'direct' per-session volume.

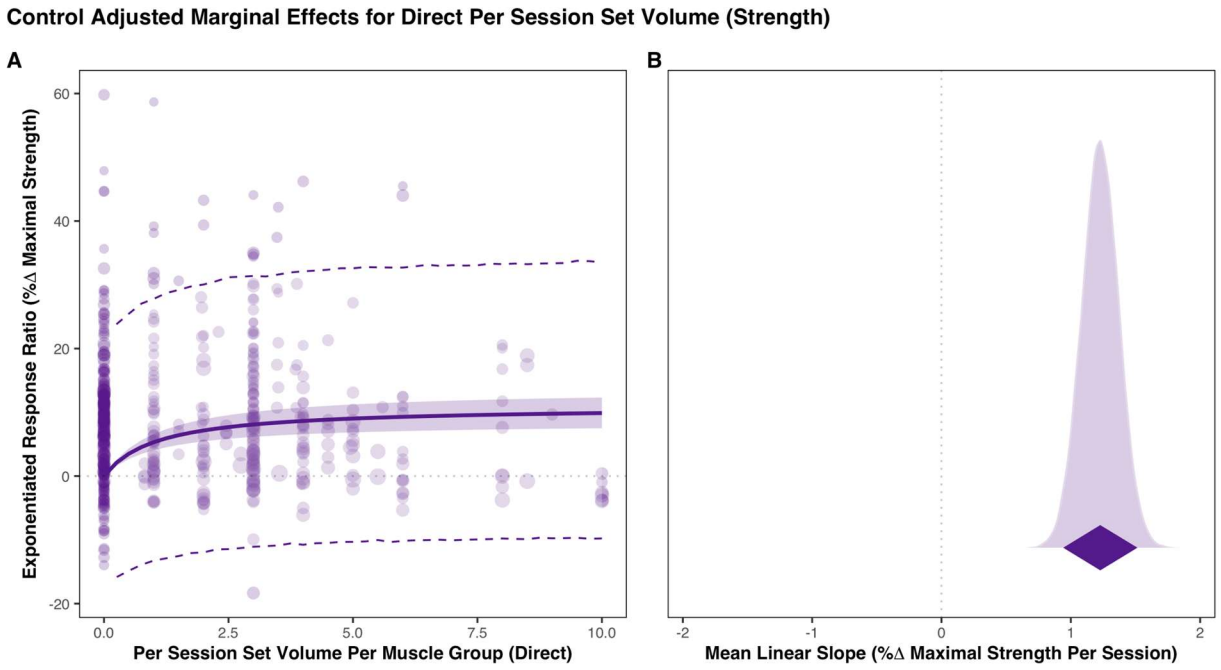

Figure 3: *Direct per-session set volume best fit multilevel meta-regression for strength (reciprocal model) analyzed as an exponentiated response ratio. Data are presented as estimated marginal means (solid line) with 95% quantile based compatibility intervals (light band = credible, dotted band = prediction) after adjusting for intervention duration and training status. Colored circles represent the effect size of each observation included in the analysis, with the size of each circle representing its weight determined by inverse variance weighting. Panel B represents the linear slope at the mean value of direct per-session volume for all effects. In all panels, the main effect for direct per-session volume is presented at the mean of the continuous fixed effect (i.e., intervention duration) and proportionally marginalized across the categorical fixed effect (i.e., training status).*

Additional models containing only direct effects (i.e., the intercept-only model comparing dichotomized higher versus lower per-session volumes and the meta-regression model of absolute per-session volume differences), including 257 total effects from 32 studies involving 972 participants, can be found in the supplementary materials (https://osf.io/7p239).

4.6 Secondary Analyses

4.6.1 Interacting Moderators

Data visualization and interpretation of interacting moderator analyses are available (https://osf.io/dqka3/). Of particular interest due to prior data suggesting a potential moderating effect are rest period length (14) and proximity to failure (49) on the dose-

response relationship between set volume and hypertrophy. Interestingly, our results do not suggest a meaningful interaction effect of either variable on strength or hypertrophy, with the uncertainty intervals of all corresponding contrasts including zero. Importantly, it should be noted that there are insufficient direct investigations into these moderators, and the number of observations contributing to these effects are often low, leading to poor estimate precision. Therefore, these moderator analyses are likely best viewed as hypothesis-generating, and we encourage extreme caution in interpretation.

4.6.2 Point of Undetectable Outcome Superiority for Per-Session Volume

Approximate values for the PUOS were determined to be 11 sets per session for 'fractional' hypertrophy (https://osf.io/3d8ts) and 2 sets per session for 'direct' strength (https://osf.io/vbzc9). It should be noted that these PUOS values are not upper limits beyond which additional increases in outcomes are not observed; rather, beyond these per-session set volume values, no pairwise comparisons to other per-session set volume values showed a >50% probability of the cumulative differences to exceed the SDES. The identified PUOS values from the favored model type were converted into the other volume quantification methods to aid interpretation (Table 1).

Table 1. Approximate conversions of PUOS from the best-fit model to other volume quantification methods

	Total sets	Direct Sets	Fractional Sets
Hypertrophy	13.835	8.165	**11***
Strength	4.957	**2***	3.478

*value from the best-fit model for that outcome, without any conversion applied. PUOS = Point of undetectable outcome superiority.

5 DISCUSSION

The present meta-regressions explored the dose-response relationships of per-session set volume on the effects of muscle hypertrophy and strength gain. The results supported our hypothesis that hypertrophy and strength gain increased as per-session volume increased, though the best-fit model indicated diminishing returns. Stronger diminishing returns were evident for strength gains compared to hypertrophy. Importantly, the PUOS for each

section can be interpreted as the threshold beyond which additional sets are unlikely to yield benefits detectable on the individual-level, probabilistically. In other words, practitioners may deem a different threshold of improvement for what is practically beneficial, potentially diverging from our smallest detectable effect size. Ultimately, it is paramount to understand that the PUOS values, while justifiable, were determined arbitrarily.

5.1 Volume Quantification Method

Previous meta-analyses exploring the effects of RT volume on muscle hypertrophy and strength gain, with the exception of the parallel project from our lab (2), either counted direct sets only (4) or did not distinguish between direct and indirect sets when accounting for weekly set volume (1,3,5), counting them equally. Importantly, treating direct and indirect sets as equivalent stimuli is unlikely to be ideal for describing the underlying data (50–53). For example, Mannarino et al. (52) observed, using a within-subjects design, significantly greater increases in elbow flexors muscle thickness with biceps curls (+11.06 ± 8.50%) than with dumbbell rows (+5.16 ± 5.61%) (SD estimated using WebPlotDigitizer) when volume was equated in men who trained to concentric muscular failure. Given the potentially lower contribution of indirect sets to muscle growth, it is unsurprising that the fractional model best explained the hypertrophy outcomes in the present study and the relationship between weekly volume and muscle growth in the parallel project from our lab (2).

Interestingly, for strength gain, the present investigation found the 'direct' set model to best describe the outcomes, while Pelland et al. (2) found the 'fractional' model to best explain the relationship between weekly set volume and strength gain. Both Pelland et al. ((2); https://osf.io/rm4xy) and the present project (https://osf.io/cuvsa) also conducted exploratory analyses to approximate the contribution of indirect sets to the volume quantification methods that were most predictive of outcomes. Regarding muscle strength, these analyses suggested that model performance across both Akaike Information Criterion and Bayesian Information Criterion was maximized quantifying indirect sets as ~39% for weekly set volume and ~16% for per-session set volume. The best-fit model for strength gain may differ from Pelland et al. (2) because the added benefit of a set on the per-session level is considerably less than on the per-week level. In Pelland et al. (2), the largest benefit for strength gain is increasing from 0 to 1 set per week, with each additional set eliciting an attenuated rate of gain; thus, increasing weekly sets from 1 to 2 confers a smaller benefit than increasing from 0 to 1. On the per-session level, the largest increase in

strength is also experienced when increasing from 0 to 1 set with diminishing returns thereafter; however, the relative importance of an increase from 1 to 2 sets is far less on a per-session level than weekly. Specifically, since the average training frequency of the included studies in this investigation for strength outcomes was 1.97 ± 0.92 sessions per week, increasing *per-session* sets from 1 to 2 would lead to an approximate increase in *weekly* sets from 2 to 4. As already outlined, a greater benefit is experienced when increasing from 1 to 2 weekly sets than from 2 to 4 weekly sets. Therefore, the increase in 1 to 2 sets per-session (~2 to 4 weekly sets) is of such small benefit that it is unsurprising that indirect sets would be of little benefit on the per-session level when even direct sets are not accounting for much additional strength gain. It should also be noted that the 'direct' quantification method has notable limitations. For instance, in studies involving isometric or isokinetic strength tests, all dynamic training was coded as indirect sets, leading to a large number of effects with solely indirect sets – in other words, 0 'direct' sets per session. A greater number of training groups with 0 sets is an inherent consequence of the 'direct' quantification method. Therefore, it is likely that indirect sets do contribute to strength gain; however, the exact degree to which those sets contribute is unclear but likely <50% of the effect of direct sets.

5.2.1 Hypertrophy Outcomes

Our meta-regressions display a positive dose-response relationship between per-session volume and hypertrophy, with diminishing returns – particularly over ~11 fractional sets – and greater uncertainty with higher per-session volumes. Greater outcomes occurred with higher per-session volumes, although these returns diminished with each additional set; above 11 'fractional' sets per session, hypertrophy outcomes continued to occur, again in a decreasing manner, although – compared to <11 sets per session – with a markedly decreased slope and greater uncertainty surrounding superiority to the outcomes at a lesser per-session volume value. Further, the relationship from zero per-session sets up until the PUOS was not linear, suggesting increasingly diminishing returns. Krieger, although unpublished with much less data (14), also identified a positive dose-response relationship. Specifically, Krieger observed diminishing returns above 6-8 'total' sets per session when interset rest was ≥2 min, but a lack of clear diminishing returns with shorter rest periods (<2 min). In contrast, we identified the PUOS at ~11 'fractional' sets per session with no meaningful effect of rest period length. These disparities may be due to differences in volume quantification method and available data to be meta-analyzed. Specifically, Krieger (14) counted 'fractional' and 'direct' sets both as one set and only included 2 studies (54,55) with ≥11 sets per session while we included 12 studies with ≥10 sets per session (54–65).

Interestingly, Benito et al. (13) identified a negative dose-response relationship whereby higher per-session volumes led to a decrease in lean body mass. This different direction of the dose-response relationship is potentially explainable by different inclusion criteria, particularly the use of indirect measures of hypertrophy (i.e., DXA) as well as different analysis techniques (i.e., not differentiating between direct and indirect sets, and analyzing the effect of per-session volume on changes in lean body mass as a secondary moderator analysis rather than a primary analysis) in Benito et al. (13).

5.2.2 Volume-Frequency Interaction: Muscle Hypertrophy

Although the present investigation only examined per-session volume, it is essential to discuss these findings in the context of weekly volume targets to understand the volume-frequency interaction. The present investigation found that the difference in hypertrophy was small and imprecise above 11 'fractional' sets (i.e., the PUOS). Pelland et al. (2) reported the PUOS occurred around 31 *weekly* 'fractional' sets. Taken together, these findings seem to suggest a frequency of 3 weekly sessions with ~10 sets per-session (10*3) to equate to the PUOS of 31 weekly sets from Pelland et al. (https://osf.io/xc3g5). However, Pelland et al. reported that the estimated effect of frequency beyond 1 was unlikely to exceed the SDES (https://osf.io/xc3g5), which seemingly contradicts the findings of the present investigation.

Multiple possibilities may explain this potential contradiction in findings. First, given the dose-response relationship of weekly volume reported by Pelland et al. (2), the effects of per-session volume should be viewed as additive throughout the week. Therefore, the most consistent comparison would be the per-session PUOS multiplied by the average weekly frequency (2.33 sessions/wk.) from the present investigation (11 * 2.33). This product of 25.6 weekly sets is closer to 31 weekly sets. This difference of ~5 weekly sets may then be explained by differences in best-fit functional forms between weekly volume (root) and per-session volume (logarithmic), limiting the direct translation between outcomes and therefore practical volume-frequency combinations. Furthermore, we encourage caution with our methods for determining the PUOS specifically as it relates to volume-frequency interactions due to inherent differences in the incremental natures of volume (i.e., a single set) and frequency (i.e., an entire session). In a relative sense, an additional increment of volume in a single session is considerably smaller in proportion compared to adding an additional session. Therefore, while the effect of frequency beyond 1 session per week in Pelland et al. (2) did not exceed the SDES, the slope was positive with a 91.3% probability of exceeding 0. Therefore, the non-zero effect of frequency and

differences in best-fit functional forms likely explain the initial contradiction, and reveal that the analyses are compatible and highlight important considerations for interpretation.

While the present data provide a direct investigation for per-session volume, insights into per-session volume-frequency interactions rely only on indirect comparisons. In other words, no study, to our knowledge, has isolated *both* of these variables to directly investigate per-session volume-frequency interactions. However, isolating both per-session volume and frequency would require four groups to investigate two levels of each variable, and even more groups would be required to control for independent effects of weekly volume. Given the intricate nature of this research question, direct investigations necessitate a resource-intensive design, warranting careful consideration before pursuing such studies.

In the absence of direct investigations, there are two competing hypotheses on the per-session volume-frequency interaction: 1) higher frequencies are only beneficial when per-session volume is low so that sufficient recovery occurs between sessions, and 2) higher frequencies are beneficial when it ensures per-session volumes do not become excessively high. For the first hypothesis, mechanistically, following the post-exercise increase, muscle protein synthesis is decreased substantially by 48 hours post-exercise (66) and possibly back to baseline in resistance-trained individuals (67). Indeed, a review by Dankel et al. (68) suggested, based upon this theory, that higher frequencies would improve muscle growth due to greater area under the curve for the protein-synthetic response; however, it's possible that most studies have a per-session volume that is too high (present study average: 5.95 ± 4.49 'fractional' sets) for this hypothesis to unfold. In other words, with a per-session volume that is too high, sufficient recovery may not occur; thus, attenuating the hypertrophic effect of another session. Interestingly, Schoenfeld et al. (69) found that trained men performing 10.5 'fractional' weekly sets for the biceps experienced greater increases in muscle thickness by performing 3.5 'fractional' sets three times per week (+6.5%) as opposed to a frequency of twice per week with 'fractional' sets split into 4.5 and 6 in each session (+4.4%) (β = 1.41; p = 0.012). Therefore, higher frequencies may confer a benefit when paired with sufficiently low per-session volumes; however, this is speculative.

Regarding the second hypothesis, it is also possible that higher frequencies are needed when per-session volumes are too high. When set volumes are high, potentially above the PUOS (i.e., 11 'fractional' sets) reported in the present investigation, there may cease to be a meaningful benefit to continuing to increase per-session volume. Furthermore, it becomes practically difficult to complete a very high number of sets in a single session;

thus, according to this hypothesis, higher frequency may enhance hypertrophy. Importantly, the present investigation does not conclusively support either hypothesis.

5.3.1 Strength Outcomes

To our knowledge this is the first published analysis of the effects of per-session volume on strength gains. The most relevant previous work is the recent parallel project from our lab (2) which meta-analyzed the effect of total weekly volume on strength gain. Although a slightly different research question, our results align with those of Pelland et al. (2) which identified a positive dose response with diminishing returns and greater uncertainty at higher weekly volumes. In the present study, higher per-session volumes lead to greater strength gains, although with a PUOS at ~2 'direct' sets per session. Greater outcomes occurred with higher per-session volumes, although with diminishing returns. Above 2 'direct' sets per session, strength outcomes continued to occur, again in a decreasing manner, although – compared to <2 sets per session – with a markedly decreased slope and greater uncertainty surrounding superiority to the outcomes at a lesser per-session volume value. This finding is in agreement with prior research showing very low volumes – as few as 1-2 sets per session with loads >80% 1RM at a rating of perceived exertion of at least 7.5 – leading to meaningful strength gains (70,71). In an attempt to aid practical interpretation, we converted the identified value of the PUOS of 2 'direct' sets per session from the favored model type into the other volume quantification methods (Table 1). Notably, the mechanisms of long-term strength development may diverge from those in the short term, therefore caution should be taken when interpreting these results.

5.3.2 Volume-Frequency Interplay: Muscle Strength

The identification of a PUOS at ~2 'direct' sets per session supplements the independent benefit of higher frequency for strength gains reported in the parallel project from our laboratory (2). Since the beneficial effects for strength are most efficient with very little per-session volume as observed in the present analysis, and there is a dose-response relationship between weekly volume and strength increases that strongly diminish above 3 'fractional' sets per week (2), a RT program consisting of low per-session volumes with a prioritization of additional frequency (up to ~2-3 sessions per week) may maximize short term strength gain based upon the available evidence.

5.4 Limitations and Considerations

This analysis has several limitations. First, while we employed wider inclusion criteria than prior analyses (1,3–10) in an effort to better investigate the potential dose-response relationships of per-session volume with hypertrophy and strength gain, it is unclear whether the present results apply to all RT scenarios. The secondary analyses provide some insight, but there is very little direct evidence regarding the impacts of these variables (i.e., participant age range, measurement methodology, standardization of load during the training intervention, etc.) on hypertrophy and strength gain and thus how different inclusion criteria may potentially have impacted the present results..

Second, the results are limited to what was explored in the included studies. For instance, not all included studies satisfactorily quantified or reported proximity to failure of all sets, resulting in the determination of failure training, nonfailure training, or a mixture rather than a specific proximity to failure that was controlled across all sets; more clarity may impact the relationship between per-session volume and hypertrophy or strength gain as previous research indicates an independent effect of proximity to failure on these outcomes (49). Furthermore, studies featuring high per-session volumes are relatively scarce. The present analysis included studies with 5.95 ± 4.49 (range = 0-24) 'fractional' sets per session for hypertrophy outcomes and 1.96 ± 2.22 (range = 0-10) 'direct' sets per session for strength outcomes; more data resulting from higher per-session volumes may alter the dose-response relationships and the functional form of the best-fit models.

Third, while great effort was taken into calculating set volumes accounting for both direct and indirect sets, it can be argued that a different quantification method could have been used in some cases. Indeed, subsequent exploratory analyses (https://osf.io/cuvsa) approximated the contribution of an indirect set to be ~32% of a direct set for hypertrophy and ~16% for strength; thus, one might reasonably critique the present study quantifying all indirect sets as 50% of a direct set. As previously discussed, this may have particularly impacted the 'direct' set model being designated as the best-fit for strength outcomes.

Fourth, most studies included in this analysis did not involve participants training with the same number of sets for all muscle groups – in fact, many studies trained only a single muscle. Therefore, while these meta-regressions illustrate the relationship between per-session volume and hypertrophy and strength, caution should be taken when attempting to extrapolate the findings to whole-body training programs. Caution should also be taken when extrapolating to larger time frames, as longer duration interventions compared to those available for the present analysis (mean duration 10.6 weeks and 10.7 weeks for

hypertrophy and strength effects, respectively) may potentially reveal different patterns to the dose-response relationships. Moreover, our analysis did not explore the potential for psychological burnout or injury risk associated with higher per-session volumes, which may become a bottleneck in practice, but may not be represented in relatively short-duration research studies.

Fifth, training-induced edema may confound post-testing muscle size measurements, particularly following protocols which elicit a large degree of muscle damage (72–74). For the studies that did report the time between the final training session and the post-testing assessments, 96.07% of the hypertrophy effects had at least 48 hours delay before measurements were taken (https://osf.io/gx2zn). However, only 69.55% of hypertrophy effects included in this analysis reported the measurement timeline, and some of those that did report it lacked precision (e.g.,, "at least 48 hours after the final training session"). Despite some data suggesting training-induced edema is unlikely to strongly influence hypertrophy measurements (75–77), this area is underexplored and therefore the present results should be interpreted with caution.

Finally, regarding strength outcomes specifically, the mean repetitions performed per set for 'direct' sets in the included studies was 9.65 ± 3.60 repetitions, with only 0.46 ± 1.23 sets of the 1.96 ± 2.22 'direct' sets performed per session being either ≥85% 1RM or ≤6 repetitions. Thus, the level of specificity for maximum strength in the included studies may not reflect the typical practices of strength athletes (e.g., powerlifters), and the results should therefore be interpreted with caution. Future research may endeavor to investigate strength gains while comparing different per-session volumes with generally heavier loads.

6 Conclusions

The relationship between per-session volume and hypertrophy is best represented by the 'fractional' set quantification method, counting direct and indirect sets as one and half a set, respectively. There is a positive dose-response relationship with additional 'fractional' per-session sets leading to greater hypertrophy, although with a PUOS at approximately 11 'fractional' sets per session and greater uncertainty at higher per-session volumes. The interaction between volume and frequency for hypertrophy is unclear and difficult to study; novel research methods may be required to gain insight. The relationship between per-session volume and strength gain is best represented by the 'direct' set quantification method, counting only sets which are the same exercise as the tested strength assessment. There is a positive dose-response relationship with more 'direct' sets leading to greater strength gain, although with a PUOS at approximately 2 'direct' sets per session and greater

uncertainty at higher per-session volumes. Based upon the results of the present study and the parallel project from our laboratory (2), a program featuring low per-session volume and an emphasis on frequency of approximately 2-3 sessions per week seems warranted to maximize short-to-moderate term strength gain.

REFERENCES

1. Baz-Valle E, Balsalobre-Fernández C, Alix-Fages C, Santos-Concejero J. A Systematic Review of The Effects of Different Resistance Training Volumes on Muscle Hypertrophy. J Hum Kinet. 2022 Jan;81:199–210.

2. Joshua Pelland, Jacob Remmert, Zac Robinson, Seth Hinson, Michael Zourdos. The Resistance Training Dose-Response: Meta-Regressions Exploring the Effects of Weekly Volume and Frequency on Muscle Hypertrophy and Strength Gain [Internet]. 2024 [cited 2025 Feb 4]. Available from: https://sportrxiv.org/index.php/server/preprint/view/460/version/587

3. Krieger JW. Single vs. Multiple Sets of Resistance Exercise for Muscle Hypertrophy: A Meta-Analysis. J Strength Cond Res. 2010 Apr;24(4):1150–9.

4. Ralston GW, Kilgore L, Wyatt FB, Baker JS. The Effect of Weekly Set Volume on Strength Gain: A Meta-Analysis. Sports Med. 2017 Dec;47(12):2585–601.

5. Schoenfeld BJ, Ogborn D, Krieger JW. Dose-response relationship between weekly resistance training volume and increases in muscle mass: A systematic review and meta-analysis. J Sports Sci. 2017 Jun 3;35(11):1073–82.

6. Currier BS, Mcleod JC, Banfield L, Beyene J, Welton NJ, D'Souza AC, et al. Resistance training prescription for muscle strength and hypertrophy in healthy adults: a systematic review and Bayesian network meta-analysis. Br J Sports Med. 2023 Sep;57(18):1211–20.

7. Cuthbert M, Haff GG, Arent SM, Ripley N, McMahon JJ, Evans M, et al. Effects of Variations in Resistance Training Frequency on Strength Development in Well-Trained Populations and Implications for In-Season Athlete Training: A Systematic Review and Meta-analysis. Sports Med. 2021 Sep;51(9):1967–82.

8. Grgic J, Schoenfeld BJ, Davies TB, Lazinica B, Krieger JW, Pedisic Z. Effect of Resistance Training Frequency on Gains in Muscular Strength: A Systematic Review and Meta-Analysis. Sports Med. 2018 May;48(5):1207–20.

9. Ralston GW, Kilgore L, Wyatt FB, Buchan D, Baker JS. Weekly Training Frequency Effects on Strength Gain: A Meta-Analysis. Sports Med - Open. 2018 Dec;4(1):36.

10. Schoenfeld BJ, Grgic J, Krieger J. How many times per week should a muscle be trained to maximize muscle hypertrophy? A systematic review and meta-analysis of studies examining the effects of resistance training frequency. J Sports Sci. 2019 Jun 3;37(11):1286–95.

11. Schoenfeld BJ, Ogborn D, Krieger JW. Effects of Resistance Training Frequency on Measures of Muscle Hypertrophy: A Systematic Review and Meta-Analysis. Sports Med Auckl NZ. 2016 Nov;46(11):1689–97.

12. Steele J, Fisher JP, Smith D, Schoenfeld BJ, Yang Y, Nakagawa S. Meta-analysis of variation in sport and exercise science: Examples of application within resistance training research. J Sports Sci. 2023 Sep 2;41(17):1617–34.

13. Benito PJ, Cupeiro R, Ramos-Campo DJ, Alcaraz PE, Rubio-Arias JÁ. A Systematic Review with Meta-Analysis of the Effect of Resistance Training on Whole-Body Muscle Growth in Healthy Adult Males. Int J Environ Res Public Health. 2020 Feb 17;17(4):1285.

14. Krieger J. Set Volume for Muscle Size: The Ultimate Evidence Based Bible [Internet]. Weightology.net; 2017. Available from: https://weightology.net/the-members-area/evidence-based-guides/set-volume-for-muscle-size-the-ultimate-evidence-based-bible/#wrap

15. Naggara O, Raymond J, Guilbert F, Roy D, Weill A, Altman DG. Analysis by Categorizing or Dichotomizing Continuous Variables Is Inadvisable: An Example from the Natural History of Unruptured Aneurysms. Am J Neuroradiol. 2011 Mar;32(3):437–40.

16. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med. 2006 Jan 15;25(1):127–41.

17. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. PLoS Med. 2009 Jul 21;6(7):e1000100.

18. Smart NA, Waldron M, Ismail H, Giallauria F, Vigorito C, Cornelissen V, et al. Validation of a new tool for the assessment of study quality and reporting in exercise training studies: TESTEX. Int J Evid Based Healthc. 2015 Mar;13(1):9–18.

19. Lin L. Graphical augmentations to sample-size-based funnel plot in meta-analysis. Res Synth Methods. 2019 Sep;10(3):376–88.

20. Nakagawa S, Lagisz M, Jennions MD, Koricheva J, Noble DWA, Parker TH, et al. Methods for testing publication bias in ecological and evolutionary meta-analyses. Methods Ecol Evol. 2022 Jan;13(1):4–21.

21. Vigotsky AD, Nuckols GL, Fisher J, Heathers J, Krieger J, Schoenfeld BJ, et al. Improbable data patterns in the work of Barbalho et al. [Internet]. SportRxiv; 2020 [cited 2025 Mar 28]. Available from: https://osf.io/sg3wm

22. Arazi H, Asadi A. Effects of 8 Weeks Equal-Volume Resistance Training with Different Workout Frequency on Maximal Strength, Endurance and Body Composition. 2011. 5(2):112–8.

23. Barcelos LC, Nunes PRP, De Souza LRMF, De Oliveira AA, Furlanetto R, Marocolo M, et al. Low-load resistance training promotes muscular adaptation regardless of vascular occlusion, load, or volume. Eur J Appl Physiol. 2015 Jul;115(7):1559–68.

24. Braith R, Graves J, Pollock M, Leggett S, Carpenter D, Colvin A. Comparison of 2 vs 3 Days/Week of Variable Resistance Training During 10- and 18-Week Programs. Int J Sports Med. 1989 Dec;10(06):450–4.

25. Carroll TJ, Abernethy PJ, Logan PA, Barber M, McEniery MT. Resistance training frequency: strength and myosin heavy chain responses to two and three bouts per week. Eur J Appl Physiol. 1998 Jul 1;78(3):270–5.

26. DeMichele PL, Pollock ML, Graves JE, Foster DN, Carpenter D, Garzarella L, et al. Isometric torso rotation strength: Effect of training frequency on its development. Arch Phys Med Rehabil. 1997 Jan;78(1):64–9.

27. Humburg H, Baars H, Schröder J, Reer R, Braumann KM. 1-Set vs. 3-Set Resistance Training: A Crossover Study. J Strength Cond Res. 2007;21(2):578.

28. Kelly SB, Brown LE, Coburn JW, Zinder SM, Gardner LM, Nguyen D. The Effect of Single Versus Multiple Setson Strength. J Strength Cond Res. 2007;21(4):1003.

29. Martín-Hernández J, Marín PJ, Menéndez H, Ferrero C, Loenneke JP, Herrero AJ. Muscular adaptations after two different volumes of blood flow-restricted training. Scand J Med Sci Sports [Internet]. 2013 Mar [cited 2025 Feb 4];23(2). Available from: https://onlinelibrary.wiley.com/doi/10.1111/sms.12036

30. McBride JM, Blaak JB, Triplett-McBride T. Effect of resistance exercise volume and complexity on EMG, strength, and regional body composition. Eur J Appl Physiol. 2003 Nov;90(5–6):626–32.

31. Naclerio F, Faigenbaum AD, Larumbe-Zabala E, Perez-Bibao T, Kang J, Ratamess NA, et al. Effects of Different Resistance Training Volumes on Strength and Power in Team Sport Athletes. J Strength Cond Res. 2013 Jul;27(7):1832–40.

32. Serra R, Saavedra FJF, Jotta B, Novaes J, Cardozo D, Alves H. The Influence Weekly Resistance Training Frequency on Strength and Body Composition. 2018. 8(1):19–24.

33. Shigaki L, Araújo CGA, Calderon MG, Costa TKC, Aguiar AF, Costa LOP, et al. Effects of Volume Training on Strength and Endurance of Back Muscles: A Randomized Controlled Trial. J Sport Rehabil. 2018 Jul 1;27(4):340–7.

34. Starkey DB, Pollock ML, Ishida Y, Welsch MA, Brechue WF, Graves JE, et al. Effect of resistance training volume on strength and muscle thickness. Med Sci Sports Exerc. 1996 Oct;28(10):1311–20.

35. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. Nature. 2019 Mar 21;567(7748):305–7.

36. Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. BMJ. 1986 Mar 15;292(6522):746–50.

37. Crippa A, Discacciati A, Bottai M, Spiegelman D, Orsini N. One-stage dose–response meta-analysis for aggregated data. Stat Methods Med Res. 2019 May;28(5):1579–96.

38. White IR, Turner RM, Karahalios A, Salanti G. A comparison of arm-based and contrast-based models for network meta-analysis. Stat Med. 2019 Nov 30;38(27):5197–213.

39. Lajeunesse MJ. On the meta-analysis of response ratios for studies with correlated and multi-group designs. Ecology. 2011 Nov;92(11):2049–55.

40. Morris SB. Estimating Effect Sizes From Pretest-Posttest-Control Group Designs. Organ Res Methods. 2008 Apr;11(2):364–86.

41. Yang Y, Williams C, Senior AM, Morrison K, Ricolfi L, Pan J, et al. Bivariate multilevel meta-analysis of log response ratio and standardized mean difference for robust and reproducible environmental and biological sciences [Internet]. Ecology; 2024 [cited 2025 Feb 4]. Available from: http://biorxiv.org/lookup/doi/10.1101/2024.05.13.594019

42. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. 1st ed. Wiley; 2019 [cited 2025 Feb 4]. Available from: https://onlinelibrary.wiley.com/doi/book/10.1002/9781119536604

43. Nakagawa S, Noble DWA, Lagisz M, Spake R, Viechtbauer W, Senior AM. A robust and readily implementable method for the meta-analysis of response ratios with and without missing standard deviations. Ecol Lett. 2023 Feb;26(2):232–44.

44. Nakagawa S, Schielzeth H. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. O'Hara RB, editor. Methods Ecol Evol. 2013 Feb;4(2):133–42.

45. Lüdecke D, Ben-Shachar M, Patil I, Waggoner P, Makowski D. performance: An R

Package for Assessment, Comparison and Testing of Statistical Models. J Open Source Softw. 2021 Apr 21;6(60):3139.

46. Lenth RV. emmeans: Estimated Marginal Means, aka Least-Squares Means [Internet]. 2017 [cited 2025 Feb 4]. p. 1.10.7. Available from: https://CRAN.R-project.org/package=emmeans

47. Pedder H, Dias S, Bennetts M, Boucher M, Welton NJ. Joining the Dots: Linking Disconnected Networks of Evidence Using Dose-Response Model-Based Network Meta-Analysis. Med Decis Making. 2021 Feb;41(2):194–208.

48. Kass RE, Raftery AE. Bayes Factors. J Am Stat Assoc. 1995 Jun;90(430):773–95.

49. Robinson ZP, Pelland JC, Remmert JF, Refalo MC, Jukic I, Steele J, et al. Exploring the Dose-Response Relationship Between Estimated Resistance Training Proximity to Failure, Strength Gain, and Muscle Hypertrophy: A Series of Meta-Regressions. Sports Med Auckl NZ. 2024 Sep;54(9):2209–31.

50. Brandão L, De Salles Painelli V, Lasevicius T, Silva-Batista C, Brendon H, Schoenfeld BJ, et al. Varying the Order of Combinations of Single- and Multi-Joint Exercises Differentially Affects Resistance Training Adaptations. J Strength Cond Res. 2020 May;34(5):1254–63.

51. Gentil P, Soares S, Bottaro M. Single vs. Multi-Joint Resistance Exercises: Effects on Muscle Strength and Hypertrophy. Asian J Sports Med [Internet]. 2015 Mar 22 [cited 2025 Feb 4];6(1). Available from: https://brieflands.com/articles/asjsm-21609.html

52. Mannarino P, Matta T, Lima J, Simão R, Freitas De Salles B. Single-Joint Exercise Results in Higher Hypertrophy of Elbow Flexors Than Multijoint Exercise. J Strength Cond Res. 2021 Oct;35(10):2677–81.

53. Schoenfeld B, Grgic J, Haun C, Itagaki T, Helms E. Calculating Set-Volume for the Limb Muscles with the Performance of Multi-Joint Exercises: Implications for Resistance Training Prescription. Sports. 2019 Jul 22;7(7):177.

54. Brigatto FA, Braz TV, Zanini TC da C, Germano MD, Aoki MS, Schoenfeld BJ, et al. Effect of Resistance Training Frequency on Neuromuscular Performance and Muscle Morphology After 8 Weeks in Trained Men. J Strength Cond Res. 2019 Aug;33(8):2104–16.

55. Zaroni RS, Brigatto FA, Schoenfeld BJ, Braz TV, Benvenutti JC, Germano MD, et al. High Resistance-Training Frequency Enhances Muscle Thickness in Resistance-Trained Men. J Strength Cond Res. 2019 Jul;33 Suppl 1:S140–51.

56. Amirthalingam T, Mavros Y, Wilson GC, Clarke JL, Mitchell L, Hackett DA. Effects of a

Modified German Volume Training Program on Muscular Hypertrophy and Strength. J Strength Cond Res. 2017 Nov;31(11):3109–19.

57. Aube D, Wadhi T, Rauch J, Anand A, Barakat C, Pearson J, et al. Progressive Resistance Training Volume: Effects on Muscle Thickness, Mass, and Strength Adaptations in Resistance-Trained Individuals. J Strength Cond Res. 2022 Mar 1;36(3):600–7.

58. Brigatto FA, Lima LE de M, Germano MD, Aoki MS, Braz TV, Lopes CR. High Resistance-Training Volume Enhances Muscle Thickness in Resistance-Trained Men. J Strength Cond Res. 2022 Jan 1;36(1):22–30.

59. Corrêa DA, Brigatto FA, Braz TV, De Carmargo JB, Aoki MS, Marchetti PH, et al. Twice-daily sessions result in a greater muscle strength and a similar muscle hypertrophy compared to once-daily session in resistance-trained men. J Sports Med Phys Fitness [Internet]. 2022 Jan [cited 2025 Feb 4];62(3). Available from: https://www.minervamedica.it/index2.php?show=R40Y2022N03A0324

60. Enes A, De Souza EO, Souza-Junior TP. Effects of Different Weekly Set Progressions on Muscular Adaptations in Trained Males: Is there a Dose-Response Effect? Med Sci Sports Exerc [Internet]. 2023 Oct 6 [cited 2025 Feb 4]; Available from: https://journals.lww.com/10.1249/MSS.0000000000003317

61. Evangelista AL, Braz TV, La Scala Teixeira CV, Rica RL, Alonso AC, Barbosa WA, et al. Split or full-body workout routine: which is best to increase muscle strength and hypertrophy? Einstein São Paulo. 2021 Aug 18;19:eAO5781.

62. Lasevicius T, Schoenfeld BJ, Grgic J, Laurentino G, Tavares LD, Tricoli V. Similar Muscular Adaptations in Resistance Training Performed Two Versus Three Days Per Week. J Hum Kinet. 2019 Aug 21;68(1):135–43.

63. Ostrowski KJ, Wilson GJ, Weatherby R, Murphy PW, Lyttle AD. The Effect of Weight Training Volume on Hormonal Output and Muscular Size and Function. J Strength Cond Res. 1997;11(3):148.

64. Pedersen H, Iversen VM, Vereide PF, Stien N, Saeterbakken AH, Fimland MS, et al. High-frequency resistance training improves maximal lower-limb strength more than low frequency. Eur J Sport Sci. 2024 May;24(5):557–65.

65. Schoenfeld BJ, Contreras B, Krieger J, Grgic J, Delcastillo K, Belliard R, et al. Resistance Training Volume Enhances Muscle Hypertrophy but Not Strength in Trained Men. Med Sci Sports Exerc. 2019 Jan;51(1):94–103.

66. Phillips SM, Tipton KD, Aarsland A, Wolf SE, Wolfe RR. Mixed muscle protein synthesis and breakdown after resistance exercise in humans. Am J Physiol-Endocrinol Metab. 1997 Jul 1;273(1):E99–107.

67. MacDougall JD, Gibala MJ, Tarnopolsky MA, MacDonald JR, Interisano SA, Yarasheski KE. The Time Course for Elevated Muscle Protein Synthesis Following Heavy Resistance Exercise. Can J Appl Physiol. 1995 Dec 1;20(4):480–6.

68. Dankel SJ, Mattocks KT, Jessee MB, Buckner SL, Mouser JG, Counts BR, et al. Frequency: The Overlooked Resistance Training Variable for Inducing Muscle Hypertrophy? Sports Med. 2017 May;47(5):799–805.

69. Schoenfeld BJ, Ratamess NA, Peterson MD, Contreras B, Tiryaki-Sonmez G. Influence of Resistance Training Frequency on Muscular Adaptations in Well-Trained Men. J Strength Cond Res. 2015 Jul;29(7):1821–9.

70. Androulakis-Korakakis P, Fisher JP, Steele J. The Minimum Effective Training Dose Required to Increase 1RM Strength in Resistance-Trained Men: A Systematic Review and Meta-Analysis. Sports Med Auckl NZ. 2020 Apr;50(4):751–65.

71. Androulakis-Korakakis P, Michalopoulos N, Fisher JP, Keogh J, Loenneke JP, Helms E, et al. The Minimum Effective Training Dose Required for 1RM Strength in Powerlifters. Front Sports Act Living. 2021;3:713655.

72. Damas F, Nosaka K, Libardi C, Chen T, Ugrinowitsch C. Susceptibility to Exercise-Induced Muscle Damage: a Cluster Analysis with a Large Sample. Int J Sports Med. 2016 Apr 26;37(08):633–40.

73. Nosaka K, Clarkson PM. Changes in indicators of inflammation after eccentric exercise of the elbow flexors: Med Amp Sci Sports Amp Exerc. 1996 Aug;28(8):953–61.

74. Nosaka K, Sakamoto K. Effect of elbow joint angle on the magnitude of muscle damage to the elbow flexors: Med Sci Sports Exerc. 2001 Jan;22–9.

75. Damas F, Phillips SM, Lixandrão ME, Vechin FC, Libardi CA, Roschel H, et al. Early resistance training-induced increases in muscle cross-sectional area are concomitant with edema-induced muscle swelling. Eur J Appl Physiol. 2016 Jan;116(1):49–56.

76. Ferreira DV, Ferreira-Júnior JB, Soares SRS, Cadore EL, Izquierdo M, Brown LE, et al. Chest Press Exercises With Different Stability Requirements Result in Similar Muscle Damage Recovery in Resistance-Trained Men. J Strength Cond Res. 2017 Jan;31(1):71–9.

77. Gordon JA, Hoffman JR, Arroyo E, Varanoske AN, Coker NA, Gepner Y, et al. Comparisons in the Recovery Response From Resistance Exercise Between Young and Middle-Aged Men. J Strength Cond Res. 2017 Dec;31(12):3454–62.