

SINGLE VERSUS MULTIPLE SETS OF RESISTANCE EXERCISE: A META-REGRESSION

JAMES W. KRIEGER

Journal of Pure Power, Colorado Springs, CO

ABSTRACT

Krieger, JW. Single versus multiple sets of resistance exercise: a meta-regression. *J Strength Cond Res* 23(6): 1890–1901, 2009—There has been considerable debate over the optimal number of sets per exercise to improve musculoskeletal strength during a resistance exercise program. The purpose of this study was to use hierarchical, random-effects meta-regression to compare the effects of single and multiple sets per exercise on dynamic strength. English-language studies comparing single with multiple sets per exercise, while controlling for other variables, were considered eligible for inclusion. The analysis comprised 92 effect sizes (ESs) nested within 30 treatment groups and 14 studies. Multiple sets were associated with a larger ES than a single set (difference = 0.26 ± 0.05 ; confidence interval [CI]: 0.15, 0.37; $p < 0.0001$). In a dose-response model, 2 to 3 sets per exercise were associated with a significantly greater ES than 1 set (difference = 0.25 ± 0.06 ; CI: 0.14, 0.37; $p = 0.0001$). There was no significant difference between 1 set per exercise and 4 to 6 sets per exercise (difference = 0.35 ± 0.25 ; CI: $-0.05, 0.74$; $p = 0.17$) or between 2 to 3 sets per exercise and 4 to 6 sets per exercise (difference = 0.09 ± 0.20 ; CI: $-0.31, 0.50$; $p = 0.64$). There were no interactions between set volume and training program duration, subject training status, or whether the upper or lower body was trained. Sensitivity analysis revealed no highly influential studies, and no evidence of publication bias was observed. In conclusion, 2 to 3 sets per exercise are associated with 46% greater strength gains than 1 set, in both trained and untrained subjects.

KEY WORDS meta-analysis, effect size, dynamic strength, volume

INTRODUCTION

Resistance exercise is an effective method for improving musculoskeletal strength, muscle mass, bone mass, and connective tissue thickness (26, 44). It is used for improving general fitness and athletic performance, for reducing injury risk, and for rehabilitation of musculoskeletal injuries (26). A resistance exercise program can also have significant health benefits, including reduction of coronary heart disease risk, improvement of glycemic control, and reduction of type II diabetes risk (26). In the elderly population, resistance exercise is an important tool for reducing osteoporosis and fracture risk, minimizing age-related sarcopenia, and improving function in activities of daily living (15, 28).

The design of a resistance exercise program requires appropriate manipulation of numerous variables, including the frequency, intensity, and volume of the program (17). For general fitness purposes, the American College of Sports Medicine has recommended a program of 8 to 10 exercises covering all major muscle groups (1). One set per exercise is recommended; each set is performed to volitional exhaustion. Each exercise is to be performed 2 to 3 days/week on nonconsecutive days. More advanced programs are recommended for athletic populations, including higher volumes of training and programmed variation in training programs (periodization) (25).

There has been considerable debate over the optimal number of sets per exercise to improve musculoskeletal strength. Some authors have contended that multiple sets are necessary to optimize strength gains, particularly in subjects with extensive resistance exercise experience (25, 45). Other authors have argued that a single set per exercise is all that is necessary, and further gains are not achieved by successive sets (11, 12, 35). In some studies, no significant difference between single and multiple sets has been observed (2, 18, 34, 43). In a qualitative review, Carpinelli and Otto (12) stated that the majority of studies comparing single with multiple sets have shown no significant differences in strength gains. However, evidence from a number of more recent studies supports greater strength gains with multiple sets (8, 16, 22–24, 31, 33, 37, 41).

These inconsistencies among studies may relate to differences in study designs or small trials with low statistical power. Thus, it may be advantageous to combine the results of

Address correspondence to James W. Krieger, jim@jopp.us.

23(6)/1890–1901

Journal of Strength and Conditioning Research

© 2009 National Strength and Conditioning Association

trials comparing single with multiple sets. In one meta-analysis of 16 studies, Rhea et al. (38) reported superior strength gains with 3-set programs as compared with 1-set programs. However, their analysis has been criticized for including studies that did not meet their reported inclusion criteria (35). Rhea et al. also included 93 multiple, non-independent effect sizes (ESs) in their analysis, which can introduce bias because of reduced variability, artificially inflated sample sizes, and an over-weighted contribution of studies contributing multiple, nonindependent data points (3). In a second meta-analysis of 140 studies and 1,433 ESs, Rhea et al. (39) concluded that 4 sets per muscle group elicited maximal strength gains in both trained and untrained subjects. However, as with their previous analysis, they included multiple, nonindependent ESs. There also were no strictly defined inclusion and exclusion criteria; as a result, the population of studies was extremely heterogeneous and included studies on ergogenic aids, diseased populations, and children. There was also no statistical analysis of the ESs and little control for study- and group-level variables that would affect the outcomes. This can make interpretation of results difficult because it cannot be clear whether the observed results represent a true-effect, a by-product of sampling variation, or a confounding effect of study- and group-level characteristics that may have correlated with set volume. Rhea et al. also analyzed their data as sets per muscle group, which cannot account for the number of sets performed of a tested exercise and the strength gains that may result from repeated efforts of that particular exercise. In a third meta-analysis, Wolfe et al. (49) examined 103 ESs from 16 studies and found multiple sets to result in significantly greater increases in strength than single sets in trained subjects ($p < 0.001$) as well as in programs lasting 17 to 40 weeks ($p < 0.05$). However, Wolfe et al. calculated ESs differently for studies with control groups vs. studies without control groups, which can introduce bias in the treatment effects (32). Also, Wolfe et al. did not control for the use of multiple exercises per target muscles; some training programs may have single sets per exercise but multiple exercises involving the same muscle groups. Finally, Wolfe et al. included studies that compared programs that varied in more characteristics than the number of sets.

The purpose of this article was to improve upon the limitations of previous meta-analyses and use hierarchical, random-effects meta-regression to compare the effects of single and multiple sets per exercise on dynamic strength. A second purpose was to establish a dose-response effect of set volume on strength. The hypothesis was that multiple sets would improve strength to a greater degree than single sets.

METHODS

Study Selection

Searches were performed of PubMed, SPORT Discus, and CINAHL for English-language studies published between January 1, 1960, and December 7, 2007. A sample of key

words and phrases used in searches included “*resistance training*,” “*strength training*,” “*resistance exercise*,” “*sets*,” “*single*,” “*multiple*,” and “*volume*.” Boolean operators such as AND, OR, and NOT were used to help narrow searches. Hand searching and cross-referencing were performed from the bibliographies of previously retrieved studies and from review articles. Studies were selected if they met the following criteria: (a) resistance exercise program lasting a minimum of 4 weeks; (b) training on at least 1 exercise for at least 1 major muscle group; major muscle groups included the quadriceps, hamstrings, pectoralis major, latissimus dorsi, biceps, triceps, and deltoids; (c) inclusion of adults 19 years of age or older; (d) comparison of single with multiple sets per exercise, with all other training variables being equivalent; (e) apparently healthy subjects, free from orthopedic limitations and comorbidities that could affect progress on a resistance exercise program; (f) pre- and post-training measurement of dynamic 1-repetition maximum (1-RM) strength; (g) sufficient data to determine sets per exercise and exercise frequency and to calculate ESs; and (h) published studies in English-language journals only.

Data Abstraction

Data were tabulated onto a spreadsheet using Microsoft Excel (Microsoft Corp., Redmond, WA). Each row represented a specific ES for a treatment group. If there were multiple ESs for a particular treatment group (i.e., a treatment group was tested on multiple exercises), then each ES was coded in a separate row.

Variables abstracted from each study were the following: authors, year, research design (randomized trial, nonrandomized trial, or randomized crossover), n , quality score, sex (male, female, or mixed), age (19–44 yr, or ≥ 45 yr), baseline body mass (kg), resistance exercise experience (< 6 mo, or ≥ 6 mo), training program duration (wk), average repetitions per set, tested exercise frequency ($\text{d} \cdot \text{wk}^{-1}$), sets per exercise, multiple exercises per target muscles (yes/no), supervised training (yes/unspecified), name of tested exercise, portion of body targeted by tested exercise (upper/lower), pre- and post-test mean 1-RM, and pre- and post-test 1-RM *SD*. A treatment group was classified as having multiple exercises per target muscles if that group performed exercises that targeted any of the prime movers of the tested exercise. For example, the prime movers of a bench press were considered to be the pectoralis major, the anterior deltoids, and the triceps. If a treatment group performed another exercise involving at least one of those muscles as a prime mover (e.g., an overhead shoulder press), then that group was considered to have performed multiple exercises per target muscles. The study quality score was the sum of 2 scores used in previous reviews to rate the quality of resistance training studies: a 0 to 10 scale-based score used by Bågenhammar and Hansson (4) and a 0 to 10 scale-based score used by Durall et al. (14). For the average repetitions per set, if a range of repetitions was reported (e.g., 8–12 repetitions), the midpoint of the range was used (e.g., 10 repetitions).

For each tested exercise in each treatment group, an ES was calculated as the pretest-post-test change in 1-RM divided by the pretest *SD* (32). Becker (5) recommended the ES for the control group be subtracted from the experimental group ES; however, numerous studies in this analysis did not include a control group. Because it is important to define the ES in a standard way across all studies (32), the control ES was assumed to be 0 in all studies and was not subtracted from the experimental ES. To test this assumption, the mean control ES was calculated among all studies that had a control group; the mean ES was -0.04 ± 0.04 , which was not significantly different from 0 ($p = 0.38$) when compared using a 1-sample *t*-test. The sampling variance for each ES was estimated according to Morris and DeShon (32). Calculation of the sampling variance required an estimate of the population ES and the pretest-post-test correlation for each individual ES. The population ES was estimated by calculating the mean ES across all studies and treatment groups (32). The pretest-post-test correlation was calculated using the following formula (32):

$$r = (S_1^2 + S_2^2 - S_D^2) / (2 S_1 S_2),$$

where s_1 and s_2 are the *SD* for the pre- and post-test means, respectively, and s_D is the *SD* of the difference scores. Where s_2 was not reported, s_1 was used in its place. Where s_D was not reported, it was estimated using the following formula:

$$S_D = \sqrt{((S_1^2/n) + (S_2^2/n))}$$

Statistical Analyses

Because multiple, nonindependent ESs were being analyzed, it was necessary to use a statistical model that accounted for this. The data presented 3 sources of variation that had to be considered: the variation between studies, the variation between treatment groups, and the variation between ESs within each treatment group (Figure 1). Thus, meta-analyses were performed using hierarchical linear mixed models, modeling the variation between studies as a random effect, the variation between treatment groups as a random effect nested within studies, the variation between ESs as a random effect nested within treatment groups, and group-level predictors as fixed effects (Figure 1) (21). This type of statistical model has been successfully used in other meta-regressions (27). The within-group variances were assumed known. Observations were weighted by the inverse of the

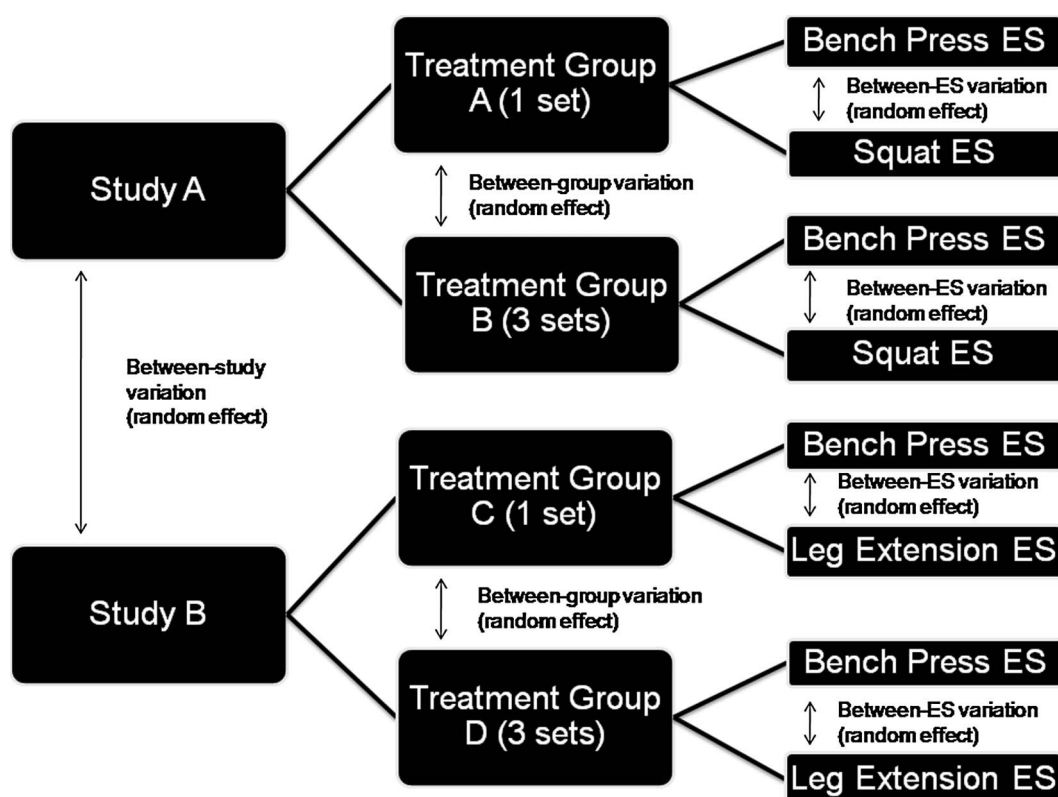


Figure 1. Random-effects, hierarchical regression model. Effect sizes (ESs) are nested within treatment groups, and treatment groups are nested within studies. Fixed-effect predictors are used to model variation between each level.

TABLE 1. Studies included in analysis.

Reference	Score†	Age, yr	Sex	Status‡	Length, wk	Supervised§	Exercise	Multiple exercises	Freq, d·wk ⁻¹	Sets	ES	SV
Arce Esquivel and Welsch (2) n = 20 n = 15	9	19–44	M,F	U	5	U	Bench press	Yes	3	1	0.41	0.01
							Leg extension	Yes	3	3	0.34	0.02
							Leg extension and chest press	Yes	3	1	1.04	0.02
Borst et al. (8) n = 11	13	19–44	M,F	U	25	Y			3	1	0.86	0.03
											0.87	0.05
n = 11										3	1.29	0.05
Galvão and Taaffe (16) n = 16 n = 16	13	≥45	M,F	U	20	Y	Bench press	Yes	2	1	0.32	0.02
							Row	Yes	2	3	0.61	0.02
							Triceps ext.	Yes	2	1	0.28	0.02
							Biceps curl	Yes	2	3	0.67	0.02
							Leg press	Yes	2	1	0.42	0.02
							Leg curl	No	2	3	0.86	0.02
							Leg extension	Yes	2	1	0.70	0.02
									2	3	1.29	0.01
									2	1	0.53	0.02
									2	3	0.67	0.02
									2	1	1.03	0.01
									2	3	1.07	0.02
									2	1	0.41	0.02
									2	3	0.91	0.02
Hass et al. (18) n = 21 n = 21	13	19–44	M,F	T	13	Y	Bench press	Yes	3	1	0.57	0.02
							Leg extension	No	3	3	0.55	0.02
							Leg curl	No	3	1	0.65	0.02
							Shoulder press	Yes	3	3	0.52	0.02
							Bicep curl	No	3	1	0.39	0.02
									3	3	0.57	0.02
									3	1	0.39	0.02
									3	3	0.61	0.02
									3	1	0.36	0.02
									3	3	0.44	0.02
Humberg et al. (22) n = 9	10	19–44	M,F	U	9	Y	Bench press	No	3	1	0.27	0.01
							Leg press (L)	No	3	3	0.38	0.01
							Leg press (R)	No	3	1	0.34	0.01
									3	3	0.64	0.01
									3	1	0.31	0.01
									3	3	0.59	0.01
									3	1	0.20	0.01
									3	3	0.44	0.01

Kemmler et al. (23) n = 29 n = 21	14	≥45	F		12	Y	Leg press	Yes	1	1	-0.02	0.01
										2	0.41	0.01
										1	-0.09	0.02
										2	0.37	0.02
							Bench press	Yes	1	1	0.00	0.01
										2	0.32	0.01
										1	-0.09	0.02
										2	0.25	0.02
							Row	Yes	1	1	0.04	0.01
										2	0.27	0.01
										1	-0.04	0.02
										2	0.27	0.02
Kraemer (experiment 2) (24) n = 20 n = 20	13	19-44	M,F	T	10	U	Bench press	Yes	3	1	0.07	0.02
							Leg press	Yes	3	3	0.19	0.02
										1	0.13	0.02
										3	0.77	0.02
McBride et al. (31) n = 9 n = 9	11	19-44	M,F	U	12	U	Leg press	No	2	1	0.58	0.06
							Bicep curl	No	2	6	1.33	0.04
										1	0.26	0.06
										6	0.59	0.07
Munn et al. (33) n = 23 n = 23 n = 23 n = 23	14	19-44	M,F	U	6	Y	Bicep curl	No	3	1	0.75	0.02
										1	0.55	0.01
										3	0.82	0.01
										3	1.13	0.02
Ostrowski et al. (34) n = 9 n = 9 n = 9	12	19-44	M	T	10	Y	Bench press	Yes	1	1	0.32	0.07
							Squat	Yes	1	2	0.48	0.07
										4	0.16	0.07
										1	0.35	0.07
Paulsen et al. (36) n = 10 n = 8	12	19-44	M	U	6	Y	Bench press	Yes	3	4	0.35	0.07
							Leg extension	Yes	3	4	0.68	0.06
							Squat	Yes	3	3	0.34	0.05
							Leg curl	No	3	1	0.23	0.09
							Shoulder press	Yes	3	1	0.34	0.04
							Lat pull	Yes	3	3	0.65	0.06
										1	0.27	0.06
										3	0.33	0.09
										1	0.25	0.05
										3	0.32	0.08
										3	0.57	0.05
										1	0.46	0.09
										3	0.39	0.05
										1	0.56	0.07

(Continued on next page)

Rhea et al. (37) n = 8 n = 8	15	19-44	M	T	12	U	Bench press	No	3	1	0.50	0.09
							Le g press	No	3	3	0.91	0.09
Rønnestad et al. (40) n = 10 n = 11	12	19-44	M	U	11	Y	Lower body	Yes	3	1	1.43	0.07
							Upper body	Yes	3	3	1.66	0.08
Schlumberger et al. (41) n = 9 n = 9	13	19-44	F	T	6	Y	Bench press	No	2	1	0.60	0.06
							Leg extension	No	2	3	1.15	0.05
										1	2.43	0.05
										3	2.83	0.06

*Freq = exercise frequency; ES = effect size; SV = sampling variation.

†Study quality score.

‡U = untrained (<6 mo resistance exercise experience); T = trained (≥6 mo experience).

§Yes (Y) or unspecified (U).

||Multiple exercises targeting same muscle groups as tested exercise (yes/no).

sampling variance (32). An intercept-only model was created, estimating the weighted mean ES across all studies and treatment groups. A full model was then created with the following predictors: quality, sex, age, resistance exercise experience, training program duration, level of supervision, targeted body portion, multiple sets per exercise (yes/no), multiple exercises per target muscles, and exercise frequency. Although the number of repetitions per set can affect strength gains (10, 46), it was not included as a predictor because of the homogeneity among studies (7-10 repetitions across all studies). The full model was then reduced by removing 1 predictor at a time, starting with the most insignificant predictor (9). The final model represented the reduced model with the lowest Bayesian information criterion (BIC) (42), and that was not significantly different ($p > 0.05$) from the full model when compared with a Likelihood Ratio Test (LRT). Model parameters were estimated by the method of restricted maximum likelihood (REML) (48); an exception occurred during the model reduction process, in which parameters were estimated by the method of maximum likelihood because LRTs cannot be used to compare nested models with REML estimates. Denominator degrees of freedom for statistical tests and confidence intervals (CIs) were calculated according to Berkey et al. (7) The multiple-sets predictor was not removed during the model reduction process. Because meta-regression can result in inflated false-positive rates when heterogeneity is present or when there are few studies (19), a permutation test described by Higgins and Thompson (19) was used to verify the significance of the predictors in the final model; 1,000 permutations were generated. Because Wolfe et al. (49) reported interactions between multiple sets per exercise and subject training experience, as well as multiple sets and training duration, 2 additional models including these interaction terms were created. Because both Paulsen et al. (36) and Rønnestad et al. (40) reported multiple sets to produce superior strength gains in the lower body but not the upper body, another model was created with the interaction between multiple sets per exercise and the half of the body targeted by the test exercise. To examine the relationship between set volume and treatment effect, a dose-response model was created by replacing the multiple-sets predictor with a categorical predictor representing the number of sets performed per exercise: 1 set, 2 to 3 sets, and 4 to 6 sets. Adjustment for post hoc multiple comparisons among set categories was performed with a Hochberg correction (20). Histograms of residuals were examined to identify major departures from normality; no departures from normality were found. Publication bias was assessed by way of a funnel plot regression method described by Macaskill et al. (29)

To identify the presence of highly influential studies that may have biased the analysis, a sensitivity analysis was carried out by removing 1 study at a time and then examining the multiple-sets predictor. Studies were identified as influential if their removal resulted in a change of greater than 1 SE in the

multiple-sets coefficient. All analyses were performed using S-PLUS version 7.0 (Insightful, Seattle, WA). Effects were considered significant at $p \leq 0.05$. Data are reported as means ($\pm SE$) and 95% CIs.

RESULTS

Study Characteristics

The analysis comprised 92 ESs nested within 30 treatment groups and 14 studies (Table 1). The weighted mean ES across all studies and treatment groups was 0.67 ± 0.11 (CI: 0.46, 0.89).

Full model. Results for the full model with all predictors are shown in Table 2. There was a significant effect of sets per exercise while controlling for all other covariates, with multiple sets being associated with a larger ES than a single set (difference = 0.26 ± 0.05 ; CI: 0.15, 0.37; $p < 0.0001$).

Reduced model. During the model reduction procedure, all covariates were found to be insignificant and were dropped from the full model. The BIC decreased from 117.0 in the full model to 81.4 in the reduced model. The reduced model was not significantly different from the full model ($p = 0.10$). The intercept for the reduced model was 0.54 ± 0.11 (CI: 0.32, 0.77; $p < 0.0001$). In the reduced model, multiple sets were associated with a larger ES than a single set (difference = 0.26 ± 0.05 ; CI: 0.15, 0.37; $p < 0.0001$) (Table 3). The mean ES for a single set was 0.54 ± 0.11 (CI: 0.32, 0.77). The mean ES for multiple sets was 0.80 ± 0.11 (CI: 0.57, 1.03).

Interactions

There was no significant interaction between multiple sets per exercise and subject training experience ($p = 0.98$), training program duration ($p = 0.34$), or the half of the body targeted by the test exercise ($p = 0.70$).

TABLE 2. Full model with all covariates.

	Coefficient $\pm$ SE*	95% confidence interval	p
Multiple sets per exercise			
No	0		
Yes	0.26 ± 0.05	(0.15, 0.37)	<0.0001
Intercept†	-3.73 ± 2.99	(-9.68, 2.22)	0.22
Age category			
19–44 yrs	1.48 ± 1.07	(-1.50, 4.46)	0.24
≥45 yrs	0		
Body portion			
Both	0		
Lower	2.09 ± 1.76	(-1.44, 5.62)	0.24
Upper	2.00 ± 1.76	(-1.53, 5.52)	0.26
Exercise frequency			
1 d/wk	0		
2 d/wk	-0.40 ± 0.78	(-2.58, 1.78)	0.64
3 d/wk	0.01 ± 0.63	(-1.74, 1.75)	0.99
Multiple exercises per target muscles			
No	0		
Yes	-0.02 ± 0.11	(-0.24, 0.19)	0.83
Quality score	0.11 ± 0.13	(-0.26, 0.48)	0.45
Sex			
F	0		
M	-1.08 ± 1.04	(-3.96, 1.80)	0.36
M, F	-1.29 ± 1.01	(-4.10, 1.52)	0.27
Training duration (wk)	0.13 ± 0.09	(-0.13, 0.39)	0.24
Training experience			
<6 mo	0		
≥6 mo	-1.05 ± 0.63	(-2.81, 0.72)	0.17
Training supervision			
Unspecified	0		
Yes	-0.20 ± 0.37	(-1.24, 0.84)	0.62

*Positive values for coefficients represent an increase in overall effect size (ES). Negative values represent a decrease in overall ES. Coefficients of 0 represent default categories in model. Coefficients for other categories within same variable represent difference from default category. †Intercept of the model produced by hierarchical regression.

TABLE 3. Sensitivity analysis.

Study removed	Coefficient*	95% confidence interval	<i>p</i>	Permutation <i>p</i>
None	0.26 ± 0.05	(0.15, 0.37)	<0.0001	<0.0001
Arce Esquivel and Welsch (2)	0.28 ± 0.05	(0.18, 0.38)	<0.0001	<0.0001
Borst et al. (8)	0.26 ± 0.05	(0.15, 0.36)	<0.0001	<0.0001
Galvão and Taaffe (16)	0.24 ± 0.06	(0.13, 0.36)	0.0001	<0.0001
Hass et al. (18)	0.29 ± 0.06	(0.17, 0.41)	<0.0001	<0.0001
Humburg et al. (22)	0.26 ± 0.06	(0.14, 0.38)	0.0001	<0.0001
Kemmler et al. (23)	0.24 ± 0.07	(0.11, 0.37)	0.0006	<0.0001
Kraemer (24)	0.25 ± 0.05	(0.14, 0.36)	<0.0001	<0.0001
McBride et al. (31)	0.25 ± 0.05	(0.14, 0.35)	<0.0001	<0.0001
Munn et al. (33)	0.25 ± 0.06	(0.14, 0.37)	<0.0001	<0.0001
Ostrowski et al. (34)	0.27 ± 0.06	(0.15, 0.38)	<0.0001	<0.0001
Paulsen et al. (36)	0.27 ± 0.06	(0.16, 0.39)	<0.0001	<0.0001
Rhea et al. (37)	0.26 ± 0.05	(0.15, 0.36)	<0.0001	<0.0001
Rønnestad et al. (40)	0.26 ± 0.05	(0.16, 0.35)	<0.0001	<0.0001
Schlumberger et al. (41)	0.25 ± 0.05	(0.14, 0.36)	<0.0001	<0.0001

*Coefficient ± SE. Value represents difference in effect size (ES) between single and multiple sets per exercise.

Dose-Response Model

In the dose-response model, 2 to 3 sets per exercise were associated with a significantly greater ES than 1 set per exercise (difference = 0.25 ± 0.06; CI: 0.14, 0.37; *p* = 0.0001).

There was no significant difference between 1 set per exercise and 4 to 6 sets per exercise (difference = 0.35 ± 0.25; CI: -0.05, 0.74; *p* = 0.17) or between 2 to 3 sets per exercise and 4 to 6 sets per exercise (difference = 0.09 ± 0.20; CI: -0.31,

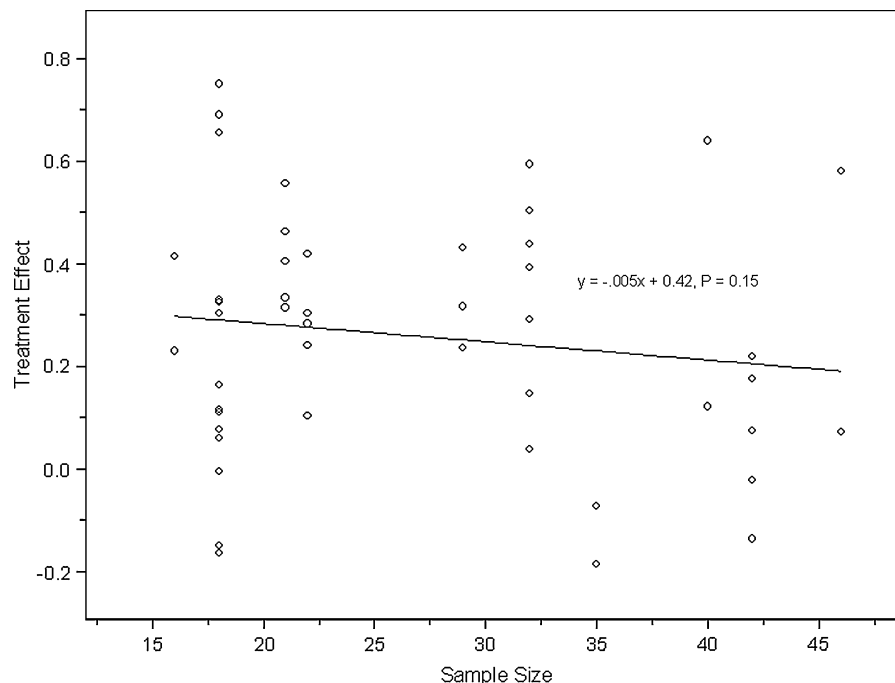


Figure 2. Funnel plot regression of sample size and treatment effect. Significant (*p* < 0.05) slope indicates potential publication bias.

0.50; $p = 0.64$). The mean ES for 1 set per exercise was 0.54 ± 0.12 (CI: 0.31, 0.77). The mean ES for 2 to 3 sets per exercise was 0.79 ± 0.12 (CI: 0.56, 1.02). The mean ES for 4 to 6 sets per exercise was 0.89 ± 0.22 (CI: 0.45, 1.32).

Sensitivity Analysis

Results for the sensitivity analysis are reported in Table 3. No influential studies were identified.

Publication Bias

A plot of treatment effect (multiple sets ES – single set ES) and sample size can be seen in Figure 2. There was no significant relationship between treatment effect and sample size (slope of line = -0.005 ± 0.004 ; $p = 0.15$), indicating no evidence of publication bias.

DISCUSSION

The purpose of this meta-analysis was to determine whether multiple sets per exercise are associated with greater strength gains than a single set per exercise in a resistance exercise program. Multiple sets per exercise were associated with significantly greater ESs in both the full and reduced statistical models. The mean ES for a single set per exercise was 0.54, whereas the mean ES for multiple sets was 0.80. Thus, multiple sets were associated with 48% greater strength gains than a single set. According to Cohen's classifications for ESs (<0.41 = small; 0.41 – 0.70 = moderate; >0.70 = large) (13), multiple sets per exercise were associated with a large treatment effect, whereas a single set was associated with a moderate treatment effect. In this dataset, the multiple-set predictor was by far the strongest predictor of strength gains because it was the only predictor remaining in the reduced model.

In 2 other meta-analyses, Rhea et al. (38) and Wolfe et al. (49) reported multiple sets to be associated with significantly greater strength gains than a single set. The current study strengthens these conclusions by improving upon the limitations of these previous works. The difference in ES between single and multiple sets in this study was similar to the difference reported by Rhea et al. (0.26 vs. 0.28, respectively); however, Rhea et al. reported a larger ES (0.70) for studies controlling all other variables other than the number of sets. Differences may relate to differences in methodology or studies included; Rhea et al. did not report which studies were classified as controlling for other variables. Rhea et al. also calculated the ES differently; they classified a single-set group as a control group, and the difference in the post-test mean between the control and multiple-set groups was used to calculate the ES. Wolfe et al. reported an ES difference of 0.41 between single and multiple sets in trained subjects only. However, Wolfe et al. defined the ES differently from the current study; they calculated the ES as the difference between the experimental group (single or multiple sets) and the control group (no training), divided by

the *SD* of the control group. For studies not including a control group, they defined the ES as the difference in pretest and post-test means divided by the pooled *SD*. They also included studies that did not control for other variables other than the number of sets.

Wolfe et al. (49) reported multiple sets to be associated with superior strength gains in trained individuals but not untrained individuals. Their results differ from the results of the current study; in the current analysis, no significant interaction between set volume and training status was observed. Thus, according to the current analysis, subjects with little resistance training experience are equally likely to benefit from multiple sets per exercise as subjects with at least 6 months of experience. The difference in this outcome between the current study and Wolfe et al. is not clear; it may relate to differences in studies included or differences in the way training status was defined (Wolfe et al. did not report how training status was defined).

Wolfe et al. (49) reported that multiple sets produced superior strength gains in training programs lasting 17 to 40 weeks but not in programs lasting for a shorter period of time. The results of the current analysis are not in agreement; there was no significant interaction between set volume and training program duration. In fact, 12 of the 14 studies in the current analysis involved studies lasting less than 17 weeks, and a significant effect of multiple sets over a single set was observed. This difference may relate to differences in studies included in the 2 analyses; only 5 of the studies included in the analysis by Wolfe et al. were included in the current analysis.

Paulsen et al. (36) and Rønnestad et al. (40) reported multiple sets to produce greater strength gains in the lower body but not the upper body. In the current analysis, there was no interaction between body portion and set volume. It is possible that these authors did not have the statistical power to detect an effect of multiple set training on the upper body; for example, in the study by Rønnestad et al., the difference in upper-body ES between 1 set and 3 sets was 0.40, but this was not statistically significant.

The dose-response analysis revealed 2 to 3 sets per exercise to be associated with 46% greater strength gains than 1 set, but no further benefit was observed with 4 to 6 sets. However, although the difference between 2 to 3 sets and 1 set was highly significant, the insignificant results for 4 to 6 sets compared with other set categories should be interpreted with caution. Only 2 of the 14 studies in this analysis consisted of treatment groups performing 4 to 6 sets per exercise (31, 34). Thus, the SEs for the estimates are large and preclude any definitive conclusions. These results differ from Rhea et al. (38), who reported a dose-response effect of an increasing number of sets up to 4 sets per muscle group, with no further benefit beyond 4 sets. However, Rhea et al. examined sets per muscle group, whereas the current analysis was on sets per exercise. Also, Rhea et al. did not statistically analyze the trend in their data or report any CIs, so it is not clear whether the

observed dose-response was a true effect or a product of sampling variation. Finally, it cannot be determined from Rhea et al. how much of the observed mean ESs were related to set volume and how much by study- and group-level related factors that were not controlled.

To examine the effects of potential outliers on the outcome, a sensitivity analysis was performed. The magnitude of the multiple-set estimate was remarkably consistent regardless of which study was removed. Thus, no particular study had a strong effect on the results.

Publication bias represents the problem in which studies showing statistically significant results are more likely to be published than studies that fail to show significant results (e.g., studies showing a significant difference between 1 set and multiple sets per exercise may be more likely to be published) (6). Thus, meta-analyses of published studies may overestimate the magnitude of a treatment effect (6). Analyses can be performed to detect the presence of publication bias; one analysis involves examining the relationship between sample size and treatment effect (29). The existence of a significant relationship suggests that publication bias may be present. However, no such relationship was observed in the current study. Wolfe et al. (49) examined the presence of publication bias using a Kendall rank correlation test and also failed to observe any evidence of bias.

The physiologic explanation behind how multiple sets are associated with superior strength gains has not been well defined. Strength training is associated with a number of structural and neural adaptations that enhance force production (26). It could be that the repetitive stimuli of multiple sets enhance hypertrophic adaptations, neural adaptations, or both. However, there will also be a point of diminishing returns. With many treatments or interventions, such as drugs, there is often an increasing response to treatment up to a certain optimal dose, and then a plateau or decrease in response beyond that dose is reached (30). The current analysis suggests that 1 set per exercise, although producing strength gains, is not the optimal dose for many individuals.

There are a number of strengths to the current study design. First, strict inclusion criteria were used; only studies comparing single with multiple sets while holding all other variables constant were included. Second, the random effects, hierarchical regression model allowed for the simultaneous modeling of the variation between studies, between treatment groups, and between ESs within each treatment group. Third, the analysis controlled for other factors that may affect strength gain, such as the presence of multiple exercises per target muscles. Fourth, both standard and permutation test *p* values were used to protect against spurious findings, a common problem with meta-regression (19). Fifth, objective quality scores were used to rate each study and were included as a covariate in the analysis to control for differences in study quality. Finally, a sensitivity analysis was

performed and indicated the results to be robust against the removal of individual studies.

One limitation to this analysis is the small number of studies incorporating 4 or more sets per exercise. Another limitation is that meta-regression, similar to epidemiologic research, can only support observational associations and cannot demonstrate causation (47). A final limitation is the availability of data (47). Some studies, despite meeting the design criteria (comparison of single versus multiple sets while keeping other variables constant), were excluded because dynamic 1-RM tests were not performed or reported. For example, Starkey et al. (43) reported no significant differences in strength gains between 1 set and 3 sets per exercise; their study was not included in this analysis because they measured isometric strength at a variety of angles rather than dynamic 1-RM strength. Because an analysis can only be undertaken for trials in which all information is available, bias can be introduced in the results (47). However, given the robustness of the results to removal of individual studies, and the lack of evidence of publication bias, it is unlikely that significant bias was present.

PRACTICAL APPLICATION

Multiple sets per exercise were associated with significantly greater strength gains than a single set per exercise during a resistance exercise program. In particular, 2 to 3 sets per exercise were associated with 46% greater gains than 1 set. This was true regardless of subject training status or training program duration. Thus, both trained and untrained subjects are equally likely to benefit from performing 2 to 3 sets per exercise. Given that the majority of the studies in this analysis compared 1 set with 3 sets, 3 sets is most likely an appropriate starting point. If time is a limiting factor, then single sets can produce strength gains, but improvements may not be optimal. No further benefits for volumes greater than 3 sets were observed, but because there were few studies involving 4 or more sets per exercise in this analysis, this conclusion should be taken with caution. Given the overwhelming evidence showing a superiority of multiple sets, further research should focus on dose-response relationships (e.g., 1 set vs. 3 sets vs. 6 sets per exercise).

ACKNOWLEDGMENTS

The author thanks Dr. Dan Wagman for his help in obtaining some articles. There were no financial or personal conflicts of interest and no external funding for this study. The results of this study do not constitute endorsement of the NSCA.

REFERENCES

1. American College of Sports Medicine Position Stand. The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. *Med Sci Sports Exerc* 30: 975–991, 1998.

2. Arce Esquivel, AA and Welsch, MA. High and low volume resistance training and vascular function. *Int J Sports Med* 28: 217–221, 2007.
3. Arthur, W JR, Bennett, Bennett, W JR, and Hufcutt, AI. *Conducting Meta-Analysis Using SAS*. Mahwah, NJ: Lawrence Erlbaum Associates, 2001.
4. Bågenhammar, S and Hansson, EE. Repeated sets or single set of resistance training: a systematic review. *Adv Physiother* 9: 154–160, 2007.
5. Becker, BJ. Synthesizing standardized mean-change measures. *Br J Math Stat Psychology* 41: 257–278, 1988.
6. Begg, CB and Berlin, JA. Publication bias and dissemination of clinical research. *J Natl Cancer Inst* 81: 107–115, 1989.
7. Berkey, CS, Hoaglin, DC, Mosteller, F, and Colditz, GA. A random-effects regression model for meta-analysis. *Stat Med* 14: 395–411, 1995.
8. Borst, SE, De Hoyos, DV, Garzarella, L, Vincent, K, Pollock, BH, Lowenthal, DT, and Pollock, ML. Effects of resistance training on insulin-like growth factor-I and IGF binding proteins. *Med Sci Sports Exerc* 33: 648–653, 2001.
9. Burnham, KP and Anderson, DR. *Model Selection and Inference: A Practical Information-Theoretic Approach*. New York, NY: Springer-Verlag, 2002.
10. Campos, GR, Luecke, TJ, Wendeln, HK, Toma, K, Hagerman, FC, Murray, TF, Ragg, KE, Ratamess, NA, Kraemer, WJ, and Staron, RS. Muscular adaptations in response to three different resistance-training regimens: specificity of repetition maximum training zones. *Eur J Appl Physiol* 88: 50–60, 2002.
11. Carpinelli, RN. Berger in retrospect: effect of varied weight training programmes on strength. *Br J Sports Med* 36: 319–324, 2002.
12. Carpinelli, RN and Otto, RM. Strength training: single versus multiple sets. *Sports Med* 26: 73–84, 1998.
13. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
14. Durall, CJ, Hermesen, D, and Demuth, C. Systematic review of single-set versus multiple-set resistance-training randomized controlled trials: implications for rehabilitation. *Crit Rev Phys Rehab Med* 18: 107–116, 2006.
15. Fielding, RA. The role of progressive resistance training and nutrition in the preservation of lean body mass in the elderly. *J Am Coll Nutr* 14: 587–594, 1995.
16. Galvão, DA and Taaffe, DR. Resistance exercise dosage in older adults: single- versus multiset effects on physical performance and body composition. *J Am Geriatr Soc* 53: 2090–2097, 2005.
17. Hass, CJ, Feigenbaum, MS, and Franklin, BA. Prescription of resistance training for healthy populations. *Sports Med* 31: 953–964, 2001.
18. Hass, CJ, Garzarella, L, De Hoyos, D, and Pollock, ML. Single versus multiple sets in long-term recreational weightlifters. *Med Sci Sports Exerc* 32: 235–242, 2000.
19. Higgins, JT and Thompson, SG. Controlling the risk of spurious findings from meta-regression. *Stat Med* 23: 1663–1682, 2004.
20. Hochberg, Y. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* 75: 800–802, 1988.
21. Hox, JJ and De Leeuw, ED. Multilevel models for meta-analysis. In: *Multilevel Modeling. Methodological Advances, Issues, and Applications*. S.P. Reise and N. Duan, eds. Mahwah, NJ: Lawrence Erlbaum Associates, 2003. pp. 90–111.
22. Humburg, H, Baars, H, Schröder, J, Reer, R, and Braumann, K-M. 1-set vs. 3-set resistance training: a crossover study. *J Strength Cond Res* 21: 578–582, 2007.
23. Kemmler, WK, Lauber, D, Engelke, K, and Weineck, J. Effects of single- vs. multiple-set resistance training on maximum strength and body composition in trained postmenopausal women. *J Strength Cond Res* 18: 689–694, 2004.
24. Kraemer, WJ. The physiological basis for strength training in American football: fact over philosophy. *J Strength Cond Res* 11: 131–142, 1997.
25. Kraemer, WJ, Adams, K, Cafarelli, E, Dudley, GA, Dooly, C, Feigenbaum, MS, Fleck, SJ, Franklin, B, Fry, AC, Hoffman, JR, Newton, RU, Potteiger, J, Stone, MH, Ratamess, NA, and Triplett-Mcbride, T. American College of Sports Medicine position stand. Progression models in resistance training for healthy adults. *Med Sci Sports Exerc* 34: 364–380, 2002.
26. Kraemer, WJ, Ratamess, NA, and French, DN. Resistance training for health and performance. *Curr Sports Med Rep* 1: 165–171, 2002.
27. Krieger, JW, Sitren, HS, Daniels, MJ, and Langkamp-Henken, B. Effect of variation in protein and carbohydrate intake on body mass and composition during energy restriction: a meta-regression. *Am J Clin Nutr* 83: 260–274, 2006.
28. Layne, JE and Nelson, ME. The effects of progressive resistance training on bone density: a review. *Med Sci Sports Exerc* 31: 25–30, 1999.
29. Macaskill, P, Walter, SD, and Irwig, I. A comparison of methods to detect publication bias in meta-analysis. *Stat Med* 20: 641–654, 2001.
30. Macdougall, J. Analysis of dose response studies: Emax model. In: *Dose Finding in Drug Development*. Ting, N, ed. New York, NY: Springer, 2006. pp. 127–145.
31. McBride, JM, Blaik, JB, and Triplett-Mcbride, T. Effect of resistance exercise volume and complexity on EMG, strength, and regional body composition. *Eur J Appl Physiol* 90: 626–632, 2003.
32. Morris, SB and Deshon, RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychology Methods* 7: 105–125, 2002.
33. Munn, J, Herbert, RD, Hancock, MJ, and Gandevia, SC. Resistance training for strength: effect of number of sets and contraction speed. *Med Sci Sports Exerc* 37: 1622–1626, 2005.
34. Ostrowski, KJ, Wilson, GJ, Weatherby, R, Murphy, PW, and Lyttle, AD. The effect of weight training volume on hormonal output and muscular size and function. *J Strength Cond Res* 11: 148–154, 1997.
35. Otto, RM and Carpinelli, RN. A critical analysis of the single versus multiple set debate. *JEPonline* 9: 32–57, 2006.
36. Paulsen, G, Mykkestad, D, and Raastad, T. The influence of volume of exercise on early adaptations to strength training. *J Strength Cond Res* 17: 115–120, 2003.
37. Rhea, MR, Alvar, BA, Ball, SD, and Burkett, LN. Three sets of weight training superior to 1 set with equal intensity for eliciting strength. *J Strength Cond Res* 16: 525–529, 2002.
38. Rhea, MR, Alvar, BA, and Burkett, LN. Single versus multiple sets for strength: a meta-analysis to address the controversy. *Res Q Exerc Sport* 73: 485–488, 2002.
39. Rhea, MR, Alvar, BA, Burkett, LN, and Ball, SD. A meta-analysis to determine the dose response for strength development. *Med Sci Sports Exerc* 35: 456–464, 2003.
40. Rønnestad, BR, Egeland, W, Kvamme, NH, Refsnes, PE, Kadi, F, and Raastad, T. Dissimilar effects of one- and three-set strength training on strength and muscle mass gains in upper and lower body in untrained subjects. *J Strength Cond Res* 21: 157–163, 2007.
41. Schlumberger, A, Stec, J, and Schmidtleicher, D. Single- vs. multiple-set strength training in women. *J Strength Cond Res* 15: 284–289, 2001.
42. Schwarz, G. Estimating the dimension of a model. *Ann Stat* 6: 461–464, 1978.
43. Starkey, DB, Pollock, ML, Ishida, Y, Welsch, MA, Brechue, WF, Graves, JE, and Feigenbaum, MS. Effect of resistance training volume on strength and muscle thickness. *Med Sci Sports Exerc* 28: 1311–1320, 1996.

44. Stone, MH. Implications for connective tissue and bone alterations resulting from resistance exercise training. *Med Sci Sports Exerc* 20: S162–S168, 1988.
45. Stone, MH, Plisk, SS, Stone, ME, Schilling, BK, O'bryant, HS, and Pierce, KC. Athletic performance development: volume load - 1-set vs. multiple sets, training velocity and training variation. *Strength Cond* 20: 22–31, 1998.
46. Stone, WJ and Coulter, SP. Strength/endurance effects from three different resistance training protocols in women. *J Strength Cond Res* 8: 231–234, 1994.
47. Thompson, SG and Higgins, JT. How should meta-regression analyses be undertaken and interpreted? *Stat Med* 21: 1559–1573, 2002.
48. Thompson, SG and Sharp, SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med* 18: 2693–2708, 1999.
49. Wolfe, BL, Lemura, LM, and Cole, PJ. Quantitative analysis of single- vs. multiple set programs in resistance training. *J Strength Cond Res* 18: 35–47, 2004.