BRIEF REVIEW

Total Number of Sets as a Training Volume Quantification Method for Muscle Hypertrophy: A Systematic Review

Eneko Baz-Valle,[1] Maelán Fontes-Villalba,[2] and Jordan Santos-Concejero[1]

[1]Department of Physical Education and Sport, University of the Basque Country UPV/EHU, Vitoria-Gasteiz, Spain; and [2]Center for Primary Health Care Research, Department of Clinical Sciences, Faculty of Medicine, Lund University, Malmö, Sweden

Downloaded from http://journals.lww.com/nsca-jscr by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdgGj2MwlZLeI= on 01/04/2024

Abstract

Baz-Valle, E, Fontes-Villalba, M, and Santos-Concejero, J. Total number of sets as a training volume quantification method for muscle hypertrophy: A systematic review. *J Strength Cond Res* 35(3): 870–878, 2021—This review aimed to determine whether assessing the total number of sets is a valid method to quantify training volume in the context of hypertrophy training. A literature search on 2 databases (PubMed and Scopus) was conducted on May 18, 2018. After analyzing 2,585 resultant articles, studies were included if they met the following criteria: (a) studies were randomized controlled trials, (b) studies compared the total number of sets, repetition range, or training frequency, (c) interventions lasted at least 6 weeks, (d) subjects had a minimum of 1 year of resistance training experience, (e) subjects' age ranged from 18 to 35 years, (f) studies reported morphologic changes through direct or indirect assessment methods, (g) studies involved subjects with no known medical conditions, and (h) studies were published in peer-reviewed journals. Fourteen studies met the inclusion criteria. According to the results of this review, the total number of sets to failure, or near to, seems to be an adequate method to quantify training volume when the repetition range lies between 6 and 20+ if all the other variables are kept constant. This approach requires further development to assess whether specific numbers of sets are key to inducing optimal muscle gains.

Key Words volume load, muscle gains, resistance training, strength training, muscle growth

Introduction

Controlling training variables is considered one of the most important factors to maximize muscle hypertrophy after resistance training (25–27,32). Previous research has reported an inverted U-shaped relationship between weekly training volume and muscle mass gains, which suggests that training volume is the most limiting variable for muscle hypertrophy (26,27). Other parameters, such as training intensity (expressed as percentage of 1 repetition maximum or %1RM), seem to be less important because similar muscle mass gains can be achieved with a wide range of repetitions when the training volume is matched (25).

The main challenge with training volume is its quantification (16). A good quantification method should imply similar muscle gains at matched training volume, independently of other training variables including intensity, training frequency, rest between sets, movement velocity, exercise order, etc. (25). Currently, the total work (TW) completed (force $[N] \cdot$ displacement $[m]$) is considered as the best quantification method when aiming for muscle hypertrophy (16). However, given the difficulty to evaluate the TW for each muscle group, other methods have been proposed as an alternative to quantify training volume. Among them, the classical volume load (VL) (sets × repetitions × kilogram) is nowadays the most used one in the scientific literature (5,12,13,29).

The main advantage of VL is that it considers many training variables (sets, repetitions, and loads), although the total number of sets may differ between different training programs matched for VL (26). This contrasts with previous research suggesting that similar muscle gains can be achieved when an equal total number of sets are completed close to muscle failure, independently of the number of repetitions and weight used (2,18,28).

Although the external load and number of repetitions (important factors for TW determination) are not taken into account when counting the total number of sets, it seems that counting the total number of sets may be an alternative

Address correspondence to Dr. Jordan Santos-Concejero, jordan.santos@ehu.eus.

35(3)/870–878

Journal of Strength and Conditioning Research
© 2018 National Strength and Conditioning Association

Copyright © 2018 National Strength and Conditioning Association. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/nsca-jscr by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdgGj2MwlZLeI= on 01/04/2024

and easier quantification method for athletes aiming at muscle hypertrophy (2,23,28).

To date, whether the total number of sets is an effective quantification method for training volume is yet to be explored. Thus, the aim of this systematic review was to analyze training programs with different arrangement of number of sets, repetition range, and training frequency to determine whether assessing the total number of sets is a valid method to quantify training volume in the context of hypertrophy training.

Methods

Procedures

A literature search was conducted on May 18, 2018. The following databases were searched: PubMed and Scopus. Databases were searched from inception up to May 2018, with no language limitation. Citations from scientific conferences were excluded.

Literature Search. The literature search was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines. In each database, the title, abstract, and keyword search fields were searched. The following keywords, combined with Boolean operators (AND/OR), were used: Muscles AND hypertrophy OR "muscle thickness" OR "cross-sectional area" AND "resistance training" AND (volume OR growth OR training). "Muscles" and "hypertrophy" are MeSH terms. No additional filters or search limitations were used. After conducting the initial search, the reference lists of articles retrieved were then screened for any additional articles that had relevance to the topic.

Inclusion Criteria. Studies were eligible for further analysis if the following inclusion criteria were met: (a) studies were randomized controlled trials comparing different groups with the number of sets explicitly reported, with the same load assignment (%RM or XRM) and without the use of external implements (i.e., pressure cuffs, hypoxic chamber, etc.), (b) studies compared the total number of sets, repetition range, or training frequency, (c) interventions lasted at least 6 weeks, (d) subjects had a minimum of 1 year of resistance training experience, (e) subjects' age ranged from 18 to 35 years, (f) studies reported morphologic changes using dual-energy x-ray absorptiometry (DXA), magnetic resonance imaging (fMRI), ultrasounds, air displacement plethysmography (ADP), or skinfolds, (g) studies involved subjects with no known medical conditions or injuries impairing training capacity, and (h) studies were published in peer-reviewed journals.

Two independent observers reviewed the studies and then individually decided whether inclusion was appropriate. In case of disagreement, a third observer was consulted. A flow chart of the search strategy and study selection is shown in Figure 1.

Study Quality. Oxford's level of evidence (20) and the Physiotherapy Evidence Database (PEDro) scale (7,14) were used by 2 independent observers to assess the methodological quality of the studies included in the systematic review. Oxford's level of evidence ranges from 1a to 5, with 1a being systematic reviews of high-quality randomized controlled trials and 5 being expert opinions. The PEDro scale consists of 11 different items related to scientific rigor. Given that the assessors are rarely blinded, and that is impossible to blind the subjects, and investigators in supervised exercise interventions, items 5–7, which are specific to blinding, were removed from the scale (25). With the removal of these items, the maximum result on the modified PEDro 8-point scale was 7 (the first item is not included in the total score) and the lowest, 0. Zero points are awarded to a study that fails to satisfy any of the included items and 7 points to a study that satisfies all the included items.

Results

Studies Selected

The search strategy yielded 2,585 studies as presented in Figure 1. Another 2 studies were added from other sources. After removing 765 duplicates, 27 studies were determined to be potentially relevant to the topic based on the information contained in the abstract, from which only 14 studies met the inclusion criteria. Excluded studies had at least one of the following characteristics: (a) subjects had little training experience or had left their training programs long time ago, or (b) the intervention period was shorter than 6 weeks, (c) the variables analyzed were not the total number of sets, repetition range, or training frequency. The overall sample for the present systematic review resulted in 359 trained athletes (352 men and 7 women) with an age range of 19–34 years (Table 1).

Level of Evidence and Quality of the Studies

According to the Oxford's level of evidence, 6 of the included 14 studies had an evidence level 1b (high-quality randomized controlled trials). The 8 remaining studies had a level of evidence 2b due to the following reasons: less than 85% of subjects completed the protocol and confidence intervals were not reported. Scores from the PEDro scale were on average 4.4 ± 0.6 and ranged from 3 to 5 (Table 2).

Evidence for Total Number of Sets as a Valid Quantification Method

Equal Number of Sets With Different Repetition Range. Six of the 14 studies comparing different repetition ranges with the same number of sets and frequency support the number of sets as an appropriate formula for quantifying training volume (2,15,18,23,24,28) (Table 3).

Schoenfeld et al. (23) compared 2 programs where a classic hypertrophy scheme (3 × 8–12RM) was used in one group and a daily undulating scheme (3 × 2–4RM/3 × 8–

Copyright © 2018 National Strength and Conditioning Association. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/nsca-jscr by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdgGj2MwlZLeI= on 01/04/2024

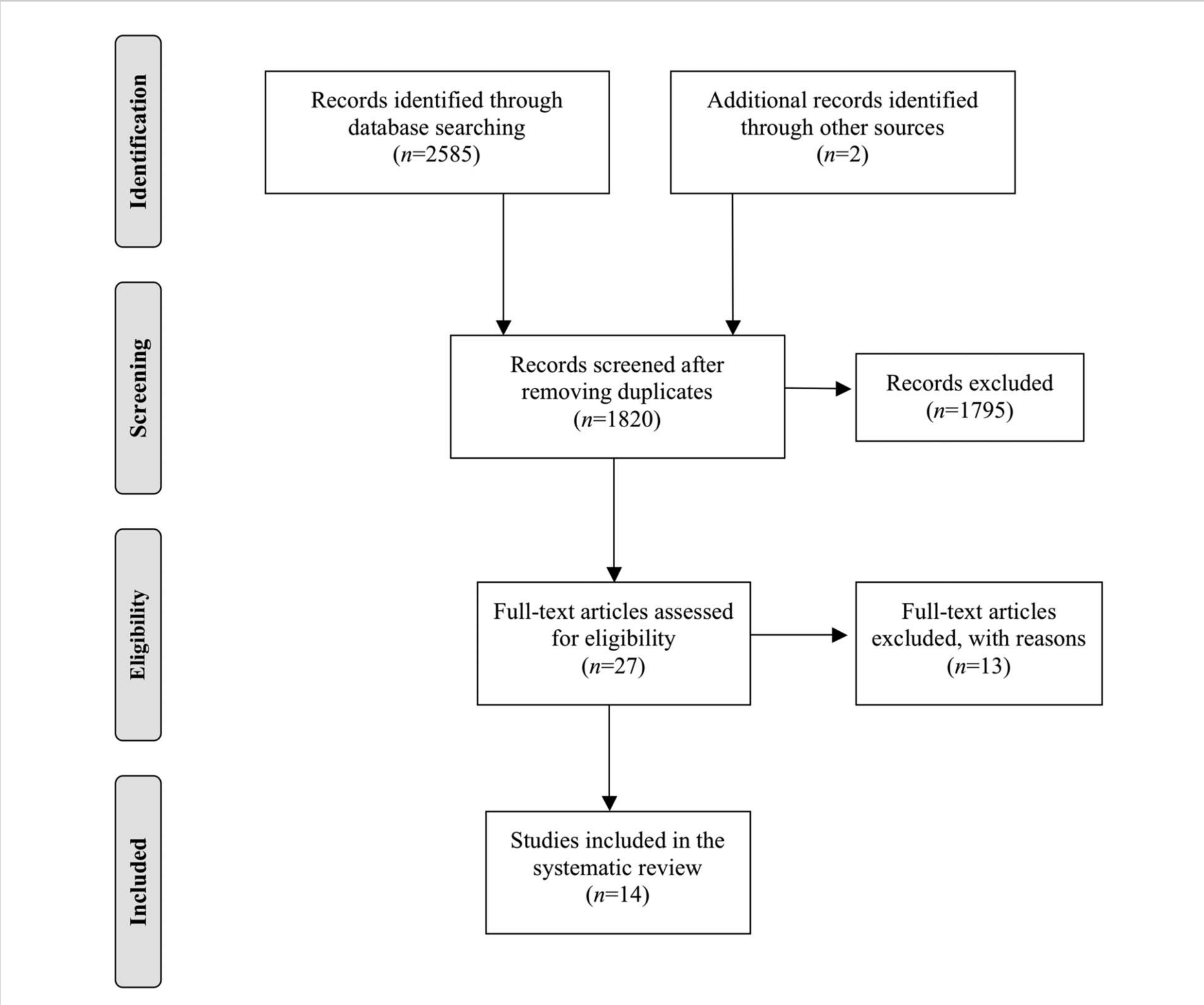

Figure 1. Flow chart of search strategy and selection of articles.

12RM/3 × 20–30RM) in the other one. Elbow flexors and extensors, and quadriceps femoris muscle thickness were measured by ultrasound. Volume load was significantly higher in the constant repetitions group, although no significant differences in muscle mass were found between groups. However, when looking at the effect sizes, the undulating program was superior (ES = 0.77 and 0.72, for extensors and flexors, respectively, moderate effect), when compared with the constant repetitions group (ES = 0.48 and 0.57, for extensors and flexors, respectively, small effect).

Another trial by the same research group (24) compared high loads (3 × 2–4RM) with moderate loads (3 × 8–12RM) training protocols during 8 weeks. They measured elbow extensors and flexors, and lateral quadriceps muscle thickness with ultrasound. There were significant differences in VL, being higher in the moderate loads group. However, no significant differences were seen for muscle thickness, although the moderate loads group induced a high effect size (ES = 1.17) compared with a low effect size in the high loads group (ES = 0.33).

Finally, Schoenfeld et al. (28) compared 2 training groups performing 3 sets per exercise with different repetition ranges. The low load group performed 25–35 repetitions to failure and the high load group 8–12 repetitions to failure in each exercise for 8 weeks. The researchers measured elbow flexors and extensors, and quadriceps femoris using ultrasound. No significant differences were observed between groups in any of the 3 measurements.

Au et al. (2) compared 3 different groups: a high repetition group (3 × 20–25RM), a low-repetitions group (3 × 8–12RM), and a control group. Subjects trained 2 times per week for 12 weeks. Fat-free mass and % body fat were determined through ADP. Results showed no significant

Copyright © 2018 National Strength and Conditioning Association. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/nsca-jscr by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdgGj2MwlZLeI= on 01/04/2024

Table 1. Characteristics of the studies and the subjects.*

Study	Number (M/F)	Age	Training experience	Other characteristics
Amirthalingam et al. (1)	19/0	19.0–24.0	At least 1 y	Subjects needed to be currently performing at least 3 training sessions per wk
Au et al. (2)	46/0	23.0 ± 2.0	At least 2 y	n/a
Brigatto et al. (4)	20/0	27.1 ± 5.5	4.1 ± 1.8 y	Minimum 1RM BS of 1.25 BM and 1RM BP of at least equal BM
Gomes et al. (9)	23/0	18.0–32.0	At least 3 y	n/a
Klemp et al. (12)	16/0	23.0 ± 3.0	At least 2 y	Minimum 1RM BS of 1.25 BM and 1RM BP of at least equal BM
Lopes et al. (13)	16/0	24.6 ± 5.8	2.3 ± 1.4 y	Subjects did not follow a periodized training program
Mangine et al. (15)	33/0	24.0 ± 3.0	5.7 ± 2.2 y	The subjects' prior training habits were under reported
Morton et al. (18)	49/0	23.0 ± 1.0	4.0 ± 2.0 y	n/a
Schoenfeld et al. (23)	30/0	23.3 ± 2.9	4.7 ± 3.2 y	Subjects regularly performed the barbell BS and BP exercises for at least 1 year prior to the study
Schoenfeld et al. (24)	26/0	23.2 ± 4.2	At least 1 y	Experienced lifters
Schoenfeld et al. (28)	24/0	23.3	3.4 y	Experienced lifters
Schoenfeld et al. (29)	20/0	23.2 ± 2.7	4.2 ± 2.4 y	Experienced lifters
Thomas and Burns (34)	12/7	34.2 ± 11.0	3.9 ± 3.8	n/a
		35.1 ± 6.9	4.6 ± 2.6 y	
Yue et al. (36)	18/0	21 ± 3.2	3.0 ± 0.5/2.9 ± 0.4 y	They included BS and BP exercises for a minimum of 2 and a maximum of 5 y
		28 ± 7.9		

*BM = body mass; BS = back squat; BP = bench press; 1RM, 1-repetition maximum.

differences between training groups, although there were significant differences when comparing the intervention groups with the control group.

Similarly, Mangine et al. (15) compared high loads (4 × 3–5RM) with moderate loads (4 × 10–12RM) during 8 weeks. Both groups before the start of the trial completed a run-in basic work period to remove confounding factors. Anthropometric variables were measured by DXA and segmental (arms, legs, and chest) muscle thickness using ultrasound. Results show no significant differences between groups in fat-free mass or segmental muscle thickness.

Morton et al. (18) conducted a trial where subjects (n = 49) trained 4 times per week during 12 weeks. One group performed a high repetition (3 × 20–25RM) and the other a low repetition program (3 × 8–12RM), and segmental muscle mass gains were measured by DXA. No significant differences were detected for fat-free mass between groups, despite lower VL in the low-repetition group.

However, since ADP and DXA poorly reflect changes in muscle mass, these results from Au et al. (2), Mangine et al. (15), and Morton et al. (18) should be interpreted with caution.

Equal Number of Sets With Different Training Frequency. Four of 14 studies analyzed the potential differences between different training frequencies matching repetition range and the number of sets per muscle group (4,9,34,36) (Table 3).

Brigatto et al. (4) compared 2 resistance training routines with different training frequencies. One of the groups trained each muscle group 1 time per week, and the other one, 2 times per week. Weekly, the number of sets per muscle group and repetition range were equalized, but total training load was not the same (different VL). The training program lasted 8 weeks and subjects trained 2 or 4 times per week. Muscle hypertrophy was assessed by muscle thickness using ultrasound (triceps braquii, elbow flexors, vastus lateralis, and

Copyright © 2018 National Strength and Conditioning Association. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/nsca-jscr by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCyw CX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdgGj2MwlZLeI= on 01/04/2024

Table 2. Physiotherapy evidence database (PEDro) ratings and oxford evidence levels of the included studies.*

Study	1	2	3	4	5	6	7	8	Total	Evidence level
Amirthalingam et al. (1)	Yes	1	0	1	1	0	1	1	5	1b
Au et al. (2)	Yes	1	0	1	0	0	1	1	4	2b
Brigatto et al. (4)	Yes	1	0	1	1	0	1	1	5	1b
Gomes et al. (9)	Yes	1	0	1	1	0	1	1	5	1b
Klemp et al. (12)	Yes	0	0	1	0	0	1	1	3	2b
Lopes et al. (13)	Yes	1	0	1	1	0	1	1	5	2b
Mangine et al. (15)	Yes	1	0	1	1	0	1	1	5	1b
Morton et al. (18)	Yes	1	0	1	1	0	1	1	5	1b
Schoenfeld et al. (23)	Yes	1	0	1	0	0	1	1	4	2b
Schoenfeld et al. (24)	Yes	1	0	1	0	0	1	1	4	2b
Schoenfeld et al. (28)	Yes	1	0	1	0	0	1	1	4	2b
Schoenfeld et al. (29)	Yes	1	0	1	0	0	1	1	4	2b
Thomas and Burns (34)	Yes	0	0	1	1	0	1	1	4	2b
Yue et al. (36)	Yes	1	0	1	1	0	1	1	5	1b
Total									4.4	

*Items in the PEDro scale: 1 = eligibility criteria were specified; 2 = subjects were randomly allocated to groups; 3 = allocation was concealed; 4 = the groups were similar at baseline regarding the most important prognostic indicators; 5 = measures of 1 key outcome were obtained from 85% of subjects initially allocated to groups; 6 = all subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least 1 key outcome were analyzed by "intention to treat"; 7 = the results of between-group statistical comparisons are reported for at least 1 key outcome; 8 = the study provides both point measures and measures of variability for at least 1 key outcome.

anterior quadriceps). No significant differences were observed between groups with regard to gains in muscle mass.

Gomes et al. (9) compared a high-frequency resistance training program with a low frequency one with the same number of sets (10–15 sets per muscle group) and repetition range (8–12 reps). The training program lasted 8 weeks and subjects trained 5 times per week. Dual-energy x-ray absorptiometry was used to calculate different body segments, and muscle mass index was estimated. No significant differences were seen between groups with regard to gains in muscle mass, regardless of significant differences in VL (higher in the high-frequency group).

Similarly, Yue et al. (36) compared 2 weekly equalized volume resistance training routines with differing training frequencies, but with the same number of sets per exercise and muscle group each week. The group with high-frequency and low-volume per session trained 4 days per week, and the group with low-frequency and high-volume per session trained 2 days per week. Both groups performed 2 different routines over the course of 6 weeks. Measurements included standard measurements for anthropometric assessment, fat mass, and fat-free mass estimated from the whole-body densitometry using ADP, limb circumference with tape, and muscle thickness (elbow flexors, anterior deltoid, and vastus lateralis) with ultrasound. No significant differences were found between groups although effect sizes were greater in the high-volume–low-frequency group (ES = 0.76 for vastus lateralis circumference and ES = 0.59 for fat-free mass).

Thomas and Burns (34) compared frequency 1 (sessions per muscle group per week) vs. frequency 3 (sessions per muscle group per week). Both groups performed 3 sets of 8–12RM in each exercise for 8 weeks. Fat-free mass and body composition were assessed by DXA. The results showed no significant differences between groups, suggesting that, independently of training frequency, there are no differences in muscle hypertrophy when training volume is matched (in terms of the number of sets and repetition range). It has to be noted that VL was not quantified.

Evidence for Volume Load as a Valid Quantification Method

Equal Volume Load With Different Number of Sets. Three of the studies included in this review support the use of VL as a strategy to quantify the training volume (12,13,29) (Table 3). Schoenfeld et al. (29) compared 2 training programs during 8 weeks with different number of sets but equal VL. One group performed a training program aiming at strength development (7 × 2–4RM), whereas the other one aimed at hypertrophy (3 × 8–12RM). The authors found no differences in biceps brachialis thickness measured with ultrasound (12.6 vs. 12.7% in the hypertrophy and strength group, respectively).

Klemp et al. (12) equated total VL and relative volume (sets × reps × %1RM) in 2 groups performing a daily undulating program with different intensities and the number of sets. The low-repetition group followed this schedule: 8 × 6 × 75% 1RM, 9 × 4 × 80% 1RM, and 10 × 2 × 85% 1RM on alternate days. Meanwhile, the high repetition group followed this schedule: 4 × 12 × 60% 1RM, 4 × 10 × 65% 1RM, and 5 × 8 × 70% 1RM on alternate days. Pectoralis major, lateral quadriceps medial, lateral quadriceps distal, and anterior quadriceps muscle thickness were measured with ultrasound. No significant differences were found between groups, although the lateral quadriceps distal

Copyright © 2018 National Strength and Conditioning Association. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/nsca-jscr by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdgGj2MwlZLeI= on 01/04/2024

Table 3. Characteristic of the training programs.*

Study	Groups	Weeks/ days	Same VL	Outcomes
Amirthalingam et al. (1)	HV: 10 × 10/60–90″ LV: 5 × 10/60–90″	6 wk, 3×	No	No significant differences between groups and similar effect sizes
Au et al. (2)	HR: 3 × 20–25RM/1′ LR: 3 × 8–12RM/1′	12 wk, 2×	Not reported	No significant differences between experimental groups, and similar effect sizes. Significant differences between experimental groups and control group
Brigatto et al. (4)	HF: each muscle group 2× LF: each muscle group 1× WV: same total sets: 8–16 sets per wk & 8–12 reps/60–120″	8 wk	No	No significant differences between groups and similar effect sizes
Gomes et al. (9)	HF: 5 × wk per muscle group HV: 1 × wk per muscle group WV: same total sets: 10–15 sets per wk & 8–12 reps/90″	8 wk	No	No significant differences between groups. Effect size favors HF group
Klemp et al. (12)	HL: 8 × 6/9 × 4/10 × 2 ML: 4 × 12/4 × 10/5 × 8 Rest: 5–7′	8 wk, 3×	Yes	No significant differences between groups and similar effect sizes
Lopes et al. (13)	ML: 6 × 10RM/60″ LL: 3 × 20RM/60″	6 wk, 4×	Yes	No significant differences between groups, and similar effect sizes
Mangine et al. (15)	VOL: 4 × 10–12 (70%)/1′ INT: 4 × 3–5 (90%)/3′	8 wk, 4×	No	No significant differences between groups in LBM, neither for segments, except in the LBM of the arm (higher effect size for the intensity group)
Morton et al. (18)	HR: 3 × 20–25RM/1′ LR: 3 × 8–12RM/1′	12 wk, 4×	No	No significant differences between groups, and similar effect sizes
Schoenfeld et al. (23)	CL: 3 × 8–12RM/2′ VAL: 3 × 2–4RM/3 × 8–12/3 × 20–30RM/2′	8 wk, 3×	No	No significant differences between groups in muscle thickness. Effect sizes favors varied group
Schoenfeld, et al. (24)	HL 3 × 2–4RM/2′ ML: 3 × 8–12RM/2′	8 wk, 3×	No	No significant differences between groups in muscle thickness. The effect size was higher in moderate load group for the vastus lateralis
Schoenfeld, et al. (28)	LL: 3 × 25–35RM/90″ HL: 3 × 8–12RM/90″	8 wk, 3×	Not reported	No significant differences between groups in all measures. The effect size was higher in high load group for the vastus lateralis
Schoenfeld et al. (29)	Strength: 7 × 2–4RM/3′ Hypertrophy: 3 × 8–12RM/90″	8 wk, 3×	Yes	No significant differences between groups and similar effect sizes
Thomas and Burns (34).	HF: each muscle group 3× LF: each muscle group 1 × 3 × 8–12/1–2′	8 wk	Not reported	No significant differences between groups and similar effect sizes
Yue et al. (36)	HF 4 $d \cdot w^{-1}$–Each muscle 2× HV 2 $d \cdot w^{-1}$–Each muscle 1× WV: same total series: 2 × 8–12/4 × 8–12 2′ rest	6 wk	Not reported	No significant differences between groups. Effect size favors high-volume group

*VL = volume load; HR = high repetition; LR = low repetition; RM, repetition maximum; HF = high frequency; HV = high volume; WV = weekly volume; HL = high loads; ML = middle loads; LL = low load; LV = low volume; VOL = volume; INT = intensity VAL; varied load; LBM = lean body mass.

Copyright © 2018 National Strength and Conditioning Association. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/nsca-jscr by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdgGj2MwlZLeI= on 01/04/2024

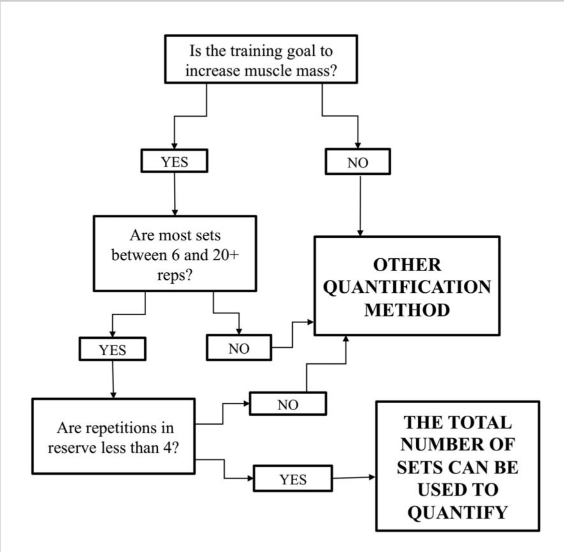

Figure 2. Proposal of training volume quantification using the total number of sets.

increased more in the high repetition group (moderate effect size) vs. the low repetition group (small effect size).

Lopes et al. (13) compared 2 training programs during 6 weeks with the same VL but different number of sets. One group performed a high set volume and medium loads program (6 × 10RM) and the other group a moderate set volume and low loads program (3 × 20RM). Changes in muscle mass were assessed by mean of skinfolds, but no differences were found between groups. Fat-free mass increased to 4.7% ± 1.0 and 3.7% ± 1.7 in the medium and low load groups, respectively. However, taking into account the lack of reliability of skinfolds when assessing muscle gains, these results should be interpreted with caution.

Evidence Against Volume Load and Total Number of Sets

One study included in this review concluded that VL and the total number of sets are not reliable strategies to quantify training volume when aiming at muscle hypertrophy (1).

Amirhalingam et al. (1), compared 2 training programs with different number of sets and different VL during 6 weeks. The high-volume group performed 10 × 10 repetitions and the low-volume group 5 × 10 repetitions. Both groups completed the same number of sets for the assistive exercises (i.e., 3–4 sets). Biceps brachii, triceps brachii, anterior thigh, and posterior thigh muscle thickness were measured with ultrasounds, and body composition was assessed with DXA. No significant differences were found between groups in muscle thickness or body composition (small effect sizes for all measures).

Discussion

The main finding of this systematic review was that the total number of sets per muscle group to muscle failure (or close to) seems to be an easy and reliable method to quantify training volume in experienced individuals who aim at hypertrophy. This systematic review also shows that other variables such as training frequency per muscle group, training intensity, and repetition range do not alter the results provided when the total number of sets is matched (2,9,15,18,23,24,34) (Table 3).

However, although most of the studies meeting the inclusion criteria supported this hypothesis, when comparing low repetition ranges (high loads) vs. moderate repetition ranges (moderate loads), the effect sizes favored training programs with moderate loads (Table 3). A recent meta-analysis comparing low vs. high loads reported that muscle mass gains were similar independently of repetition range and external load (25). Yet, in most studies, the authors described the low repetition protocol as a range of 6–12 repetitions (25). For these cases, the present review suggests that using the total number of sets to quantify training volume is more accurate when the repetition range lies between 6 and 20+ repetitions than when it lies below 6 repetitions.

Among the different mechanisms leading to muscle hypertrophy, metabolic stress is one of the most important factors (31). Metabolic stress involves cell swelling, changes in myokine production, generation of reactive oxygen species, and an acute hormonal response (31), although this last point may be more related to exercise adaptations (35). Metabolic stress induces muscle adaptations mainly by increasing motor unit recruitment (6), which implies that when using low loads, it is key to reach muscle failure (or at least get close to it) (30).

This suggests that when comparing different training protocols or quantifying training sessions with a wide range of repetitions, sets must be close to muscle failure (3 or less repetitions in reserve) (22). However, although performing sets to failure can potentiate muscle mass gains, doing it in a systematic way may be detrimental at a neuromechanic (force generation capacity) and hormonal levels (cortisol and testosterone alterations) (11,17). Therefore, to optimize training quantification using the number of sets, it is recommended to take into account sets (close) to muscle failure.

However, well-designed experiments still support the use of VL to quantify training volume (5). Campos et al. (5) reported that low and intermediate repetition training protocols induce similar muscular adaptations and suggested that both physical performance and the associated physiological adaptations are linked to the intensity and number of repetitions performed, thus supporting the "strength-endurance continuum." Yet, the aforementioned study only tested untrained subjects. Training responses have been reported to be dependent of training experience because experienced subjects have an attenuated postexercise anabolic response (10,19,21). This implies that the results reported by Campos et al. (5) may not be extrapolated to trained athletes as the ones featured in the review.

Copyright © 2018 National Strength and Conditioning Association. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/nsca-jscr by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdgGj2MwlZLeI= on 01/04/2024

Another major constraint of VL as a quantification method becomes evident when different exercises are compared. For example, if we use VL to compare leg press and squats at an equivalent relative intensity and using the same number of sets and repetitions (to failure), total tonnage and thus, TW, will always be higher for leg press. However, in a real-life scenario, the perceived effort when performing these exercises will be similar, or even higher for the squat, because it requires more muscles to complete the movement (because of the stabilization requirements) and because of a higher muscle activation (8). Therefore, when the goal is to compare training loads at different times of season and the recommended exercises vary, then VL may not be the best choice for training volume quantification (16,33). In addition, according to Amirthalingam et al. (1), when certain training volume is achieved, there are no further muscle gains. This may explain why studies supporting VL as a quantification method reported similar results independently of the total number of sets (29).

We have to acknowledge several limitations. These include the moderate quality of some the studies analyzed. According to the PEDro scale used in this review, only the half of the studies included scored 5 points (Table 2), whereas 6 studies scored 4 points and 1 study scored 3 points, which ultimately precludes us from drawing a definitive conclusion from the results. Another important limitation was related to the methods used to assess the outcomes. For example, one of the studies (13) used skinfolds as an assessment tool and, although it is validated (3), is one of the less reliable methods. The same is true for the ADP, which was used by one of the studies included in this review (2). Finally, the results of this review cannot be extrapolated to the general population because it only analyzed trained subjects. All these limitations imply that the conclusions of this review should be interpreted with caution.

In summary, although there are many variables in a training program that may influence muscle mass gains, it seems that counting the number of sets to failure, or near to, is an adequate method to quantify training volume when all the other variables are kept constant. Therefore, and based on the results discussed above, we suggest the approach depicted in Figure 2 for an optimal training volume quantification using the total number of sets.

Practical Applications

The results of this review suggest that counting the total number of sets to, or near to, failure per muscle group can be an optimal strategy to quantify training volume in experienced athletes aiming at hypertrophy. More specifically, the total number of sets can be used when the repetition range lies between 6 and 20+ (Figure 2). Athletes and coaches can use this method to monitor the increasing load throughout a training mesocycle, as well as to compare training loads between different training blocks in an easy and reliable way. This approach requires further development to assess whether specific numbers of sets are key to inducing optimal muscle gains and may help to guide experimental enquiries into hypertrophy focused training along a slightly different path than currently being tread.

Acknowledgments

There is no disclosure of funding to report for this study. Authors report no conflict of interest.

References

1. Amirthalingam, T, Mavros, Y, Wilson, G, Clarke, J, Mitchell, L, and Hackett, D. Effects of a modified German volume training program on muscular hypertrophy and strength. *J Strength Cond Res* 31: 3109–3119, 2017.
2. Au, J, Oikawa, S, Morton, R, Macdonald, M, and Phillips, S. Arterial stiffness is reduced regardless of resistance training load in young men. *Med Sci Sports Exerc* 49: 342–348, 2017.
3. Bentzur, K, Kravitz, L, and Lockner, D. Evaluation of the BOD POD for estimating percent body fat in collegiate track and field female athletes: A comparison of four methods. *J Strength Cond Res* 22: 1985–1991, 2008.
4. Brigatto, F, Braz, T, Zanini, T, Germano, M, Aoki, M, Schoenfeld, B, et al. Effect of resistance training frequency on neuromuscular performance and muscle morphology after eight weeks in trained men. *J Strength Cond Res* 33: 2104–2116, 2019.
5. Campos, G, Luecke, T, Wendeln, H, Toma, K, Hagerman, F, Murray, T, et al. Muscular adaptations in response to three different resistance-training regimens: Specificity of repetition maximum training zones. *Eur J Appl Physiol* 88: 50–60, 2002.
6. Dankel, S, Mattocks, K, Jessee, M, BucknerS, Mouser, J, and Loenneke, J. Do metabolites that are produced during resistance exercise enhance muscle hypertrophy? *Eur J Appl Physiol* 117: 2125–2135, 2017.
7. de Morton, N. The PEDro scale is a valid measure of the methodological quality of clinical trials: A demographic study. *Aust J Physiother* 55: 129–133, 2009.
8. Escamilla, R, Fleisig, G, Zheng, N, Lander, J, Barrentine, S, Andrews, J, et al. Effects of technique variations on knee biomechanics during the squat and leg press. *Med Sci Sports Exerc* 33: 1552–1566, 2001.
9. Gomes, G, Franco, C, Nunes, P, and Orsatti, F. High-frequency resistance training is not more effective than low-frequency resistance training in increasing muscle mass and strength in well-trained men. *J Strength Cond Res* 33: S130–S13, 2019.
10. Gonzalez, A, Hoffman, J, Townsend, J, Jajtner, A, Wells, A, Beyer, K, et al. Association between myosin heavy chain protein isoforms and intramuscular anabolic signaling following resistance exercise in trained men. *Physiol Rep* 3: e12268, 2015.
11. Izquierdo, M, Ibañez, J, González-Badillo, J, Häkkinen, K, Ratamess, N, Kraemer, W, et al. Differential effects of strength training leading to failure versus not to failure on hormonal responses, strength, and muscle power gains. *J Appl Physiol* 100: 1647–1656, 2006.
12. Klemp, A, Dolan, C, Quiles, J, Blanco, R, Zoeller, R, Graves, B, et al. Volume-equated high- and low-repetition daily undulating programming strategies produce similar hypertrophy and strength adaptations. *Appl Physiol Nutr Metab* 41: 699–705, 2016.
13. Lopes, C, Aoki, M, Crisp, A, de Mattos, R, Lins, M, da Mota, G, et al. The effect of different resistance training load schemes on strength and body composition in trained men. *J Hum Kinet* 58: 177–186, 2017.
14. Maher, C, Sherrington, C, Herbert, R, Moseley, A, and Elkins, M. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther* 83: 713–721, 2003.
15. Mangine, G, Hoffman, J, Gonzalez, A, Townsend, J, Wells, A, Jajtner, A, et al. The effect of training volume and intensity on improvements in muscular strength and size in resistance-trained men. *Phys Rep* 3: e12472, 2015.

Copyright © 2018 National Strength and Conditioning Association. Unauthorized reproduction of this article is prohibited.

16. McBride, J, McCaulley, G, Cormie, P, Nuzzo, J, Cavill, M, and Triplett, N. Comparison of methods to quantify volume during resistance exercise. *J Strength Cond Res* 23: 106–110, 2009.

17. Morán-Navarro, R, Pérez, C, Mora-Rodríguez, R, de la Cruz-Sánchez, E, González-Badillo, J, Sánchez-Medina, L, et al. Time course of recovery following resistance training leading or not to failure. *Eur J Appl Physiol* 117: 2387–2399, 2017.

18. Morton, R, Oikawa, S, Wavell, C, Mazara, N, McGlory, C, Quadrilatero, J, et al. Neither load nor systemic hormones determine resistance training-mediated hypertrophy or strength gains in resistance-trained young men. *J Appl Physiol* 121: 129–138, 2016.

19. Nader, G, von Walden, F, Liu, C, Lindvall, J, Gutmann, L, Pistilli, E, et al. Resistance exercise training modulates acute gene expression during human skeletal muscle hypertrophy. *J Appl Physiol* 116: 693–702, 2014.

20. Oxford Centre for Evidence-based Medicine. *Levels of Evidence*. Oxford, UK: University of Oxford, 2009. pp. 4–5.

21. Phillips, S, Tipton, K, Ferrando, A, and Wolfe, R. Resistance training reduces the acute exercise-induced increase in muscle protein turnover. *Am J Physiol Endocrinol Metab* 276: 118–124, 1999.

22. Sampson, J and Groeller, H. Is repetition failure critical for the development of muscle hypertrophy and strength? *Scand J Med Sci Sports* 26: 375–383, 2015.

23. Schoenfeld, B, Contreras, B, Ogborn, D, Galpin, A, Krieger, J, and Sonmez, T. Effects of varied versus constant loading zones on muscular adaptations in trained men. *Int J Sports Med* 37: 442–447, 2016.

24. Schoenfeld, B, Contreras, B, Vigotsky, A, and Peterson, M. Differential effects of heavy versus moderate loads on measures of strength and hypertrophy in resistance-trained men. *J Sports Sci Med* 15: 715–722, 2016.

25. Schoenfeld, B, Grgic, J, Ogborn, D, and Krieger, J. Strength and hypertrophy adaptations between low- vs. high-load resistance training: A systematic review and meta-analysis. *J Strength Cond Res* 31: 3508–3523, 2017.

26. Schoenfeld, B and Grgic, J. Evidence-based guidelines for resistance training volume to maximize muscle hypertrophy. *J Strength Cond Res* 40: 107–112, 2018.

27. Schoenfeld, B, Ogborn, D, and Krieger, J. Dose-response relationship between weekly resistance training volume and increases in muscle mass: A systematic review and meta-analysis. *J Sports Sci* 35: 1073–1082, 2017.

28. Schoenfeld, B, Peterson, M, Ogborn, D, Contreras, B, and Sonmez, G. Effects of low-vs. high-load resistance training on muscle strength and hypertrophy in well-trained men. *J Strength Cond Res* 29: 2954–2963, 2015.

29. Schoenfeld, B, Ratamess, N, Peterson, M, Contreras, B, Sonmez, G, and Alvar, B. Effects of different volume-equated resistance training loading strategies on muscular adaptations in well-trained men. *J Strength Cond Res* 28: 2909–2918, 2014.

30. Schoenfeld, B. Is there a minimum intensity threshold for resistance training-induced hypertrophic adaptations? *Sports Med* 43: 1279–1288, 2013.

31. Schoenfeld, B. Potential mechanisms for a role of metabolic stress in hypertrophic adaptations to resistance training. *Sports Med* 43: 179–194, 2013.

32. Schoenfeld, B. *Science and Development of Muscle Hypertrophy*. Champaign, IL: Hum Kinet, 2016.

33. Scott, B, Duthie, G, Thornton, H, and Dascombe, B. Training monitoring for resistance exercise: Theory and applications. *Sports Med* 46: 687–698, 2016.

34. Thomas, M and Burns, S. Increasing lean mass and strength: A comparison of high frequency strength training to lower frequency strength training. *Int J Exer Sci* 9: 159, 2016.

35. Walker, S, Häkkinen, K, Haff, GG, Blazevich, A, and Newton, R. Acute elevations in serum hormones are attenuated after chronic training with traditional isoinertial but not accentuated eccentric loads in strength-trained men. *Phys Rep* 5: e13241, 2017.

36. Yue, F, Karsten, B, Larumbe-Zabala, E, Seijo, M, and Naclerio, F. Comparison of 2 weekly-equalized volume resistance-training routines using different frequencies on body composition and performance in trained males. *Appl Physiol Nutr Metab* 43: 475–481, 2018.

Downloaded from http://journals.lww.com/nsca-jscr by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdgGj2MwlZLeI= on 01/04/2024

Copyright © 2018 National Strength and Conditioning Association. Unauthorized reproduction of this article is prohibited.