Effects of resistance training with moderate vs heavy loads on muscle mass and strength in the elderly: A meta-analysis

R. Csapo¹, L. M. Alegre²

¹Institute of Sport Science, University of Innsbruck, Innsbruck, Austria, ²GENUD Toledo Research Group, University of Castilla-La Mancha, Toledo, Spain

Corresponding author: Luis M. Alegre, PhD, GENDU Toledo Research Group, Faculty of Sports Sciences, University of Castilla-La Mancha, Avda. Carlos III s/n, 45071 Toledo, Spain. Tel: +34 925268800, ext. 5506, Fax: +34 925268846, E-mail: luis.alegre@uclm.es

Accepted for publication 12 July 2015

The purpose of the present study was to perform a meta-analysis to compare the efficacy of heavy (∼80% of one repetition maximum, 1RM) vs light-moderate load (∼45% 1RM) resistance training (RT) programs in inducing strength gains and skeletal muscle hypertrophy in elderly people. To assess the role of training volumes, studies in which training protocols were matched for mechanical work were independently analyzed. In all 15 studies included (448 subjects, age 67.8 years), when comparing heavy with light-moderate loads, strength gains tended to be larger following RT with higher intensities of load, with the resulting total population effect being μ = 0.430 (P = 0.060). Effect sizes were substantially smaller in “work-matched” studies (μ = 0.297, P = 0.003). Training with higher loads also provoked marginally larger gains in muscle size, although the degree of training-induced muscle hypertrophy was generally small (0.056 < μ < 0.136). To conclude, provided a sufficient number of repetitions is performed, RT at lower than traditionally recommended intensities of load may suffice to induce substantial gains in muscle strength in elderly cohorts.

It is widely accepted that resistance training (RT) promotes increases in skeletal muscle mass and strength. As research over the past 20 years has convincingly demonstrated, RT is particularly important from the 5th–6th decade of life onwards, since it currently represents the only effective and widely applicable tool to control and even revert sarcopenia, i.e., the age-associated losses in muscle size and function (Hakkinen et al., 1998; Reeves et al., 2004; Morse et al., 2005a, b, 2007; Narici & Maganaris, 2007). Thereby, RT contributes significantly to improved mobility (Symons et al., 2005; Krist et al., 2013) as well as enhanced quality of life and overall health in elderly populations (Kell et al., 2001; Rejeski & Mihalko, 2001; Weening-Dijkstra et al., 2011). To provoke these desirable training adaptations, world-leading organizations in exercise-related research (ACSM, 2009b; Peterson & Gordon, 2011) have recommended that RT be performed at relatively high intensities of load, which should be gradually increased from 60% to 70% to over 80% of the individual one repetition maximum (1RM). In support of these recommendations, several meta-analyses comparing the effectiveness of RT as performed at different intensities have confirmed that training at higher loads is associated with relatively greater gains in muscle size and strength (Peterson et al., 2010; Steib et al., 2010; Silva et al., 2014).

Despite the undeniable efficacy of RT with heavy loads, in recent years, controversy has arisen about the question whether strength and muscle mass gains would also be achievable with lighter load RT programs (Schuenke et al., 2012, 2013; Burd et al., 2013; Schoenfeld, 2013a). Several researchers have challenged the notion that only heavy loads might serve as the exclusive driver of RT-associated muscular adaptations. Their main criticism relates to the fact that studies comparing different RT regimen often failed to control for differences in the total mechanical work performed or the degree of fatigue induced by the training intervention (Fisher et al., 2011; Burd et al., 2012; Raymond et al., 2013). In fact, there is evidence that, when matched for mechanical work, RT provokes substantial gains in muscle mass and strength, irrespective of whether heavy or rather lighter loads were used (Léger et al., 2006; Alegre et al., 2015). Similar observations were made in studies investigating the effectiveness of low-load RT assisted by artificial blood flow restriction to invoke early muscular fatigue (Loenneke et al., 2012; Schoenfeld, 2013a). Consequentially, the results of previous meta-analyses comparing the efficacy of different resistance training intensities (Peterson et al., 2010; Steib et al., 2010; Silva et al., 2014) may have been biased by the inclusion of studies that failed to control for unequal amounts of mechanical work or degrees of
training-induced fatigue. At the least in previously untrained individuals, training with lower loads might represent a potent (Schoenfeld, 2013a) and possibly even equally effective stimulus of muscular adaptations (Raymond et al., 2013) as conventional, heavy-load RT.

Senior subjects represent a cohort that might particularly benefit from the possibility to lower RT loads. This is because the use of heavy loads may be contraindicated in subjects suffering from uncontrolled hypertension or cardiovascular disease (Williams et al., 2007; Thompson et al., 2013) – ailments that are particularly common in this age group. Also, training with heavier loads has been related to a higher rate of perceived exertion, even when the total training load was carefully matched (Alegre et al., 2015). Coincident or possibly consequential to the greater sensation of effort, research not only suggests that older adults prefer lower exercise intensity (King et al., 1991), but also that an inverse relationship between exercise adherence and intensity exists (Perri et al., 2002).

The goals of the present manuscript are, therefore, to perform a meticulous review of articles comparing the efficacy of heavy and light-moderate load RT as opposed to no exercise in elderly cohorts. More specifically, we aimed to extract the results reflecting RT-induced changes in muscle size and strength for the statistical analysis of pooled data. Unlike previous meta-analyses of data obtained in elderly people (Peterson et al., 2010; Steib et al., 2010; Raymond et al., 2013), we independently analyzed the effects of work-matched RT on strength and muscle hypertrophy.

**Methods**

To retrieve the articles for this meta-analysis, the online databases PUBMED and MEDLINE were systematically searched for the following combinations of terms: (a) “resistance training” or “strength training,” combined with “intensity” or “load” and “hypertrophy”; (b) “resistance training” or “strength training,” combined with “moderate intensity,” “low intensity,” “low load” or “moderate load”; and (c) “resistance training” or “strength training,” combined with “intensity” or “load” and “hypertrophy” or “cross-sectional area.” Screening the abstracts of the resulting list of articles, studies to compare the effects of light or moderate RT to conventional heavy-load RT programs, as recommended by the American College of Sports Medicine (ACSM, 2009a, b), were selected. High-intensity RT was defined as training where loads were progressively increased to 80% of 1RM or higher, whereas interventions using average maximum loads of 60% of 1RM or lower were considered as low-moderate intensity RT. In one study (Kerr et al., 1996), training intensities were not reported in terms of data obtained in elderly people (Peterson et al., 2010; Raymond et al., 2013), we independently analyzed the effects of work-matched RT on strength and muscle hypertrophy.

**Methods**

To retrieve the articles for this meta-analysis, the online databases PUBMED and MEDLINE were systematically searched for the following combinations of terms: (a) “resistance training” or “strength training,” combined with “intensity” or “load” and “hypertrophy”; (b) “resistance training” or “strength training,” combined with “moderate intensity,” “low intensity,” “low load” or “moderate load”; and (c) “resistance training” or “strength training,” combined with “intensity” or “load” and “hypertrophy” or “cross-sectional area.” Screening the abstracts of the resulting list of articles, studies to compare the effects of light or moderate RT to conventional heavy-load RT programs, as recommended by the American College of Sports Medicine (ACSM, 2009a, b), were selected. High-intensity RT was defined as training where loads were progressively increased to 80% of 1RM or higher, whereas interventions using average maximum loads of 60% of 1RM or lower were considered as low-moderate intensity RT. In one study (Kerr et al., 1996), training intensities were not reported in terms of the percentages of 1RM but rather set based on multiple (8RM vs 20RM) repetition maxima; according to Bryzcki’s equation (Bryzcki, 1993) these loads coincide with approximately 80% and 47% of 1RM, respectively. Several studies compared the outcomes of RT as performed at three different intensities of load. Where outcomes did not differ significantly between groups, the results of the two groups following the most similar training prescriptions were merged by calculating the pooled mean and standard deviation [merged: 70% and 90% to 80% 1RM (Beneka et al., 2005); 47% and 63% to 50% 1RM (Fatiouos et al., 2006); 75% and 84% to 80% 1RM (Harris et al., 2004)], to facilitate the calculation of total population effects. In cases where merging of results was not feasible because of significant between-group differences, only the results reported for the group training at the higher of the two lower loads were extracted [results from groups training at 20% 1RM discarded in (Van Roie et al., 2013)]. To be included, studies had to provide data reflecting either training-induced changes in muscle size or strength. Imaging-based measures of muscle cross-sectional area or volume, analyses of body composition by either dual-energy X-ray absorptiometry (DXA) or whole-body air displacement plethysmography, and anthropometric techniques were accepted as measures of muscle size. Muscle strength had to be evaluated by either determination of 1RM or isokinetic or isometric dynamometry. To minimize the number of potentially biasing factors, movement velocity, i.e., the time to complete a repetition, had to be similar in both the groups training at high and low-moderate intensities (Pereira & Gomes, 2003). Since movement velocities were expected to differ between groups, we therefore excluded a study in which power-oriented resistance training was performed at either 20%, 50% or 80% of 1RM (de Vos et al., 2005). Further, only studies performed in humans aged 50 years or older were included. This age limit was defined because it typically coincides with the onset (Faulkner et al., 2007) or aggravation (Janssen et al., 2000) of age-associated loss in skeletal muscle mass and function. Only papers that were published in peer-reviewed journals and available in English language were considered for further analysis. As opposed to a previous meta-analysis (Raymond et al., 2013), we excluded studies that reported secondary treatment approaches, such as administration of steroids, hormone replacement therapy or the intake of any further substances that could affect the responsiveness to resistance training (Sullivan et al., 2005; 2007). It should be noted that failure to report the intake of these substances in included studies represents a potential source of bias. As opposed to studies involving the administration of anabolic substances, however, we included one study where subjects were administered nutritional supplementation (Onambélé-Pearson et al., 2010). Studies on the effects of RT with blood flow restriction were also excluded. Both literature research and selection were independently performed by both authors and agreement was reached on the final list of studies to be included. The evidence on which this review is based is up-to-date as of May 2015.

In cases where descriptive statistics were separately reported for both training groups (high vs low load group) and points of time (pre- and post-training), group means were extracted and used to calculate the raw between-group differences in changes of means D (Eqs.[1]):

$$D = (G_{post} - G_{pre}) - (G_{post} - G_{pre})$$

[1]

where G1 and G2 are the mean values reported for group 1 and 2 before (pre) and after (post) the training intervention. The respective standard deviations $SD_{G1/2, post}$ were used to calculate the group specific standard deviation of change $SD_{G1/2}$ (Eqs.[2]):

$$SD_{G1/2} = \sqrt{SD_{G1/2, post}^2 + SD_{G1/2, post}^2}$$

[2]

Since the correlations between pre- and post-training measures were typically not reported, we used a conservative estimate ($r_{pre-post} = 0.7$) as proposed by Khoury et al. (2013). The pooled standard deviation of change $SD_{pooled}$ was then determined under consideration of the sample sizes $N_{G1/2}$ in both groups (Eqs.[3]):

$$SD_{pooled} = \sqrt{((N_{G1} - 1) \cdot SD_{G1}^2 + (N_{G2} - 1) \cdot SD_{G2}^2) \cdot (N_{G1} + N_{G2} - 2)^{-1}}$$

[3]

Subsequently, Cohen’s d and its standard error $SE_{d}$ were calculated for all included studies according to Eqs.[4] and [5]:

$$d = \frac{mean_{G2} - mean_{G1}}{SD_{pooled}}$$

$$SE_{d} = \frac{SD_{pooled}}{\sqrt{N_{G1} + N_{G2}}}$$

[4] [5]
\[ d = D \cdot SD_{pooled}^{-1} \]  \hfill (4)

\[ SE_r = \sqrt{1 \cdot (n_{g1}^{-1} + n_{g2}^{-1} + d^2 \cdot (2 \cdot n_{g1} + n_{g2})^{-1})^{-1}} \]  \hfill (5)

Finally, for ease of interpretation, \( d \) and \( SE_r \) were converted to the scale of Pearson’s \( r \), which serves as the measure of effect size in the present meta-analysis (Eqns.[6] and [7]):

\[ r = d \cdot \sqrt{d^2 + 4}^{-1} \]  \hfill (6)

\[ SE_r = \sqrt{16 \cdot SE_r^2 \cdot ((d^2 + 4)^{-1})^{-1}} \]  \hfill (7)

Several studies directly reported means and standard deviations of the changes in the target variables, requiring only the calculation of the pooled standard deviation and effect sizes as detailed in Eqns.[3–7]. Commonly, results from multiple related outcome measures, i.e., data of muscle strength or size from different muscle groups, were reported. On such occasions, to avoid a unit-of-analysis error, a composite effect size \( \overline{r} \) was calculated as the mean of the \( n \) single effect sizes. The variance of \( \overline{r} \) is given by

\[ \text{var}_{\overline{r}} = \frac{1}{m^2} \left( \sum_{i=1}^{m} r_i + \sum_{i=1}^{m} \left( \frac{n_i \sqrt{\text{var}_{r_i}} \sqrt{\text{var}_{r_i}}}{\text{var}_{r_i}} \right) \right) \]  \hfill (8)

where \( r_i \) is the correlation between single outcomes. Since this was typically not known, this factor was conservatively assumed to be \( r_i = 1 \), which may have led to an overestimation of variance and, thus, an underestimation of precision. In controlled studies, independent analyses were performed to compare high and low load as well as control group against each other.

Since the studies incorporated in this meta-analysis differed both in terms of the precise training parameters as well as the age and sex of the trained individuals, a random effects model was used to estimate the grand population effect \( \mu \). Such models assume that the intervention effects observed in different studies are normally distributed around \( \mu \) with a variance of \( \tau^2 \), which was estimated based on the observed effect sizes \( r \), and the associated variances \( \text{var}_r \), using the DerSimonian and Laird method (DerSimonian & Laird, 1986). Effect sizes were then weighted by the inverse of their respective variances, which was calculated as the sum of \( \text{var}_r \) and \( \tau^2 \). The total population effect \( \mu \) was then estimated as the sum of the weighted effect sizes \( r_i \), divided by the sum of weights. The variance of the total population effect, calculated as the reciprocal of the sum of weights, was used to compute confidence intervals for \( \mu \). A \( z \)-value to test the null hypothesis that \( \mu = 0 \) was computed by dividing \( \mu \) by its standard error. Finally, a forest plot was generated to visually compare the single effect sizes with the total population effect. Funnel plots were created to graphically investigate the possibility of publication bias and other sources of effect size heterogeneity, and Egger’s regression tests (Egger et al., 1997) were carried out as additional tests of funnel plot asymmetry. The above statistical procedures were performed in agreement with the recommendations by Borenstein et al. (2009) and carried out using commercially available software (Comprehensive Meta Analysis Version 2.2, Biostat Inc., Englewood, New Jersey, USA).

For all included studies and training groups, the total amount of physical work performed in each session was calculated as the product of sets \( \times \) repetitions \( \times \) load and used to classify training regimens as either “work-matched” or “non-matched.” In addition, the Physiotherapy Evidence Database (PEDro) score (Maher et al., 2003) was used to assess the methodological quality of all included studies. Each article was screened using this tool by both authors, and consensus was reached regarding the final scores of individual studies.

Results

The initial literature research resulted in a total of 1919 articles. After discarding of duplicates, the titles and abstracts of 868 articles were screened. Fifty-five of these papers were retrieved for full text analysis, out of which 15 matched the criteria for inclusion. It should be noted that two studies by Fatouros et al. (2005b, 2006) were carried out in the same cohort. Consequently, the results reflecting training-induced changes in muscle strength were merged for the calculation of one composite effect size. The flowchart depicted in Fig. 1 reflects the precise process of literature research and selection.

Out of the 15 studies finally included, 12 (Pruitt et al., 1995; Bemben et al., 2000; Hortobágyi et al., 2001; Vincent et al., 2002; Kalapotharakos et al., 2004; Seynnes et al., 2004; Beneka et al., 2005; Fatouros et al., 2005a,b, 2006; Singh et al., 2005; Cassilhas et al., 2007) were performed as randomized controlled trials, and one used as quasi-controlled design in which only one hemisphere was trained and the contralateral side served as the non-exercising control (Kerr et al., 1996), although data reflecting changes in muscle strength on the non-trained side were not reported. The remaining two studies were uncontrolled (Onambélé-Pearson et al., 2010; Van Roie et al., 2013), although allocation to training groups still occurred on a randomized basis. The level of evidence was relatively homogeneous across studies, with PEDro scores consistently ranging between 5 and 7 points (maximum: 10 points). Higher scores were hindered by the imminent difficulties to conceal the group allocation process (0/15 studies), as well as to blind subjects (2/15 studies), therapists (0/15) and assessors (1/15). The detailed results of the PEDro analyses are evident from Table 1.

Across all studies, a total of 448 subjects (202 females, 246 males) were trained at either high (\( n = 230 \), 98 females; age: 67.8 \( \pm \) 7.3 years) or lower (\( n = 218 \), 95 females; age: 67.9 \( \pm \) 7.0 years) intensity of load. The total amount of physical work performed in each training session was matched in 11 studies (73%). The training duration ranged from 56 to 365 days, with an average of 154 \( \pm \) 100 days. In all but four works that focused exclusively on the knee extensor muscles (Hortobágyi et al., 2001; Seynnes et al., 2004; Beneka et al., 2005; Fatouros et al., 2005a,b, 2006; Van Roie et al., 2013), the training interventions included multiple exercises to target all major muscle groups. Training was performed three times per week in all studies included. Across all studies, subjects performed 21.3 \( \pm \) 5.7 and 37.9 \( \pm \) 18.0 repetitions with the higher (80.8 \( \pm \) 2.0% 1RM) and lower intensity of load (44.4 \( \pm \) 9.9% 1RM), respectively. Considering only the 11 studies that were matched for physical work, between-group differences in the total number of repetitions performed were significant (20.6 \( \pm \) 6.0 vs 43.8 \( \pm \) 18.0 repetitions, \( P = 0.001 \)), whereas repetition numbers were equal in non work-matched studies.
Changes in muscle strength

Data on training-induced changes in muscle strength were reported in all 15 studies included. All studies favored resistance training with higher intensities of load, with effect sizes $r$ ranging, however, between 0.038 (Van Roie et al., 2013) and 0.990 (Singh et al., 2005). The resulting total population effect was calculated to be $\mu = 0.430$, with the effect just failing to reach statistical significance ($P = 0.060$; see forest plot in Fig. 2(a)). A funnel plot relating the studies’ effect sizes to the inverse of their standard error (i.e., their influence on the total population effect) demonstrated that a disproportionate number of studies reported effect sizes that were smaller than the total population effect. As is evident from Fig. 2(b), one particularly powerful study (Singh et al., 2005), characterized by the lowest standard error, was found to lie far to the right of the 95% confidence region. A significant Egger’s test ($t = 2.294,$ $P = 0.041$) provided further evidence for funnel plot asymmetry. These findings suggest that our comparison of high vs low-moderate RT in terms of strength gains may have been affected by an outlier, resulting in bias towards greater beneficial effects of high-intensity training. Indeed, after exclusion of the study by Singh et al. (2005), both Funnel plot (not shown) and Egger’s test ($t = 0.947,$ $P = 0.364$) reflected symmetry. To assess whether results might differ in dependency of whether the training programs were matched for physical work, an independent analysis was run only for the 11 studies classified as “work-matched.” The so determined population effect was found to be $\mu = 0.297$ (CI: 0.102–0.471, $P = 0.003$). As compared with non-training control groups, both training interventions provoked strong and significant gains in muscle strength (high loads: $\mu = 0.778$, CI: 0.447–0.921, $P < 0.001$; low loads: $\mu = 0.663$, CI: 0.396–0.826, $P < 0.001$). Average
increases in strength in the “work-matched” studies were 43% and 35% (high and light-moderate loads, respectively).

Changes in muscle size
Training-induced changes in muscle dimensions were reported in a total of seven studies. Measures included ultrasound-based measurements of muscle thickness (Onambélé-Pearson et al., 2010) or cross-sectional area (Bemben et al., 2000), computed tomography (Kalapotharakos et al., 2004; Van Roie et al., 2013), analyses of body composition by DXA (Vincent et al., 2002) or whole-body air displacement plethysmography (Cassilhas et al., 2007) as well as the anthropometric determination of thigh circumference (Fatouros et al., 2005a). Just as for the changes in muscle strength, all studies found the training with higher intensities of load to be more effective in provoking muscle hypertrophy. Composite effect sizes \( r \) ranged between 0.008 (Fatouros et al., 2005a) and 0.309 (Cassilhas et al., 2007). The total population effect was estimated to be \( \mu = 0.136 \) \((P = 0.036; \text{see forest plot in Fig. 3(a)})\). The Funnel plot (Fig. 3(b)) and associated Egger’s test \( (t = 0.151, P = 0.618) \) provided neither visual nor numerical evidence of publication bias. After exclusion of two studies characterized as “non-matched,” this metric decreased to \( \mu = 0.056 \) \((-0.098–0.207, P = 0.480)\). Importantly, comparison with control group data reported in five studies (Bemben et al., 2000; Vincent et al., 2002; Kalapotharakos et al., 2004; Fatouros et al., 2005a; Cassilhas et al., 2007) revealed that the overall gains in muscle size following training with both high \( (\mu = 0.199, CI: 0.046–0.343, P = 0.011) \) and low intensities of load \( (\mu = 0.108, CI: -0.050–0.261, P = 0.179) \) were small and, for the lower loads, non-significant. Average increases in muscle size in the “work-matched” studies that assessed appendicular muscle mass (Bemben et al., 2000; Kalapotharakos et al., 2004; Fatouros et al., 2005a; Van Roie et al., 2013) were 11% and 9% (high and light-moderate loads, respectively).

Discussion
The goal of the present meta-analysis was to compare the efficacy of heavy vs light-moderate load RT programs in increasing skeletal muscle mass and strength in elderly cohorts. Summarizing the results of 15 original articles, we found that, as compared with non-training controls, both high- and lighter load training programs may induce significant gains in strength \( (0.659 < \mu < 0.769) \). Total population effects further suggest that training effects may be more pronounced when using heavier loads \( (\mu = 0.430) \), although between-group differences were substantially weaker in studies where training protocols were matched for the total amount of physical
| Study | Work-\(\text{matched/}\text{controlled\)(\(n\text{ and mean age})\}\text{)} \text{Subjects} \text{Duration in} \text{Training} \text{Training} \text{Outcome measures} \text{Body part trained and assessed} |
|---|---|---|---|---|---|
| Bemben et al. (2000) | Yes \(\text{yes}\) | \(n\) HIT: 10 (50.5 years); \(n\) LIT: 7 (51.9 years); | 180 (72) | HIT: 3 × 3 × 80%; LIT: 3 × 16 × 40% | Same as outcome measures | Elbow extension, elbow flexion, hip abduction, hip adduction, hip flexion, hip extension, knee extension, knee flexion, latissimus pull, leg press, seated row, shoulder press | LE, UE |
| Bemeka et al. (2005) | Yes \(\text{yes}\) | \(n\) HIT: 32 (69.4 years); \(n\) LIT: 16 (69.4 years); | 112 (48) | HIT: 3 × 6–8 × 80%; LIT: 3 × 12–14 × 50% | Leg curls, leg extension, leg press | Knee extension (isokinetic) | LE |
| Cassilhas et al. (2007) | No \(\text{yes}\) | \(n\) HIT: 20 (68.4 years); \(n\) LIT: 19 (69.0 years); | 168 (72) | HIT: 2 × 8 × 80%; LIT: 2 × 8 × 50% | Same as outcome measures | Abdominal crunch, chest press, knee flexion, knee extension, lower back extension, vertical traction | CHEST PRESS, LEG PRESS | LE, UE, T |
| Fatouros et al. (2005a) | Yes \(\text{yes}\) | \(n\) HIT 20 (72.4 years); \(n\) LIT (70.3 years) | 168 (72) | HIT: 2–3 × 6–8 × 80–85%; LIT: 2–3 × 14–16 × 50–55% | Chest press, knee extension, knee flexion, leg press, shoulder press, latissimus pull, elbow flexion, elbow extension | CHEST PRESS, LEG PRESS | CHEST PRESS, LEG PRESS | LE, UE |
| Fatouros et al. (2005b) | Yes \(\text{yes}\) | \(n\) HIT: 14 (70.8 years); \(n\) LIT: 26 (70.5 years) | 168 (72) | HIT: 2–3 × 8 × 80%; LIT: 2–3 × 14–16 × 55% | Chest press, knee extension, knee flexion, leg press, shoulder press, latissimus pull, elbow flexion, elbow extension, abdominal crunch, lower back extension | CHEST PRESS, LEG PRESS | CHEST PRESS, LEG PRESS | LE, UE |
| Hortobágyi et al. (2001) | Yes \(\text{yes}\) | \(n\) HIT: 9 (72.8 years); \(n\) LIT: 9 (73.1 years); | 70 (30) | HIT: 5 × 4–6 × 80%; LIT: 5 × 8–12 × 40% | Knee extension, knee flexion, chest press, latissimus pull, elbow flexion, elbow extension, abdominal crunch, lower back extension | Knee extension (concentric, eccentric, isometric) | LE |
| Kalapotharakos et al. (2004) | Yes \(\text{yes}\) | \(n\) HIT: 11 (64.6 years); \(n\) LIT: 12 (65.7 years); | 84 (36) | HIT: 3 × 8 × 89%; LIT: 3 × 15 × 60% | Knee extension, knee flexion, chest press, latissimus pull, elbow flexion, elbow extension, abdominal crunch, lower back extension | CHEST PRESS, LEG PRESS | CHEST PRESS, LEG PRESS | LE, UE |
| Kerr et al. (1996) | Yes \(\text{yes}\) | \(n\) HIT: 23 (58.4 years); \(n\) LIT: 19 (55.7 years); | 365 (156) | HIT: 3 × 8 × 89%; LIT: 3 × 20 × 50% | Same as outcome measures | Knee extension, leg press, ankle rotation, hip extension, additional theraband exercises | Ankle plantarflexion, hip extension, leg extension (isometric and isoinertial), leg press | LE |
| Onambélé-Pearson et al. (2010) | No \(\text{no}\) | \(n\) HIT: 13 (68.1 years); \(n\) LIT: 18 (71.8 years); | 84 (36) | HIT: 2–4 × 8–11 × 80%; LIT: 2–4 × 8–11 × 40% | Knee extension, knee flexion, leg press, bench press, latissimus pull, shoulder press, elbow flexion, hip abduction, hip adduction, lower back extension | Knee flexion, bench press, shoulder press, latissimus pull, lower back extension, hip abduction, hip adduction, knee extension, knee flexion, leg press | LE, UE |
| Pruitt et al. (1995) | Yes \(\text{yes}\) | \(n\) HIT: 8 (67.0 years); \(n\) LIT: 7 (67.6 years); | 365 (156) | HIT: 2 × 7 × 80%; LIT: 2 × 14 × 40% | Knee extension, chest press, bench press, latissimus pull, shoulder press, elbow flexion, hip abduction, hip adduction, lower back extension | Elbow flexion, bench press, shoulder press, latissimus pull, lower back extension, hip abduction, hip adduction, knee extension, knee flexion, leg press | LE, UE |
| van Roie et al. (2013) | Yes \(\text{yes}\) | \(n\) HIT: 18 (67.7 years); \(n\) LIT: 19 (67.4 years); | 84 (36) | HIT: 2 × 12–16 × 80%; LIT: 60 × 20% × 20 × 40% | Knee extension, leg press | Knee extension (isometric, isokinetic and isoinertial), leg press | CHEST PRESS, LEG PRESS | CHEST PRESS, LEG PRESS | LE, UE |
| Seynnes et al. (2004) | No \(\text{yes}\) | \(n\) HIT: 8 (83.3 years); \(n\) LIT: 6 (80.7 years); | 70 (30) | HIT: 3 × 8 × 80%; LIT: 3 × 8 × 40% | Knee extension, chest press, knee flexion, leg press, chest press, shoulder press, seated row | Knee extension (isometric, isokinetic and isoinertial), leg press | CHEST PRESS, LEG PRESS | CHEST PRESS, LEG PRESS | LE, UE |
| Singh et al. (2005) | No \(\text{yes}\) | \(n\) HIT: 20 (69.0 years); \(n\) LIT: 20 (70.0 years); | 56 (24) | HIT: 3 × 8 × 80%; LIT: 3 × 8 × 20% | Knee extension, chest press, knee flexion, leg press, calf press, chest press, seated row, shoulder press, hip adduction, lower back extension, hip abduction, hip adduction, lower back extension | CHEST PRESS, SHOULD PRESS | CHEST PRESS, SHOULD PRESS | LE, UE, T |
| Vincent et al. (2002) | Yes \(\text{yes}\) | \(n\) HIT: 24 (66.6 years); \(n\) LIT: 22 (67.6 years); | 168 (72) | HIT: 1 × 8 × 80%; LIT: 1 × 13 × 50% | Knee extension, knee flexion, leg press, chest press, shoulder press, seated row | CHEST PRESS, SHOULD PRESS | CHEST PRESS, SHOULD PRESS | LE, UE, T |
| Summary | 73% \(\text{yes}\)/ 87% | \(\text{yes}\) | \(n\) HIT: 230 (67.8 years); \(n\) LIT 218 (67.9 years) | 153 (66) | | | | |

**HIT, high-intensity resistance training; LE, lower extremity; LIT, light-moderate intensity resistance training; T, trunk; UE, upper extremity.**
work performed ($\mu = 0.297$). Analysis of the subset of studies to report data on the training-related changes in muscle size revealed that neither the higher nor lower load RT programs performed were effective in inducing significant muscle hypertrophy ($0.108 < \mu < 0.199$).

Conventional training prescriptions as published by world leading organizations in exercise-related research suggest that older adults engage at least twice a week in muscle strengthening activities at moderate to vigorous intensity (ACSM, 2009a; Peterson & Gordon, 2011) to prevent loss of muscle mass and function and, thus, counter sarcopenia (Narici & Maganaris, 2007). Such intensities coincide with resistances equivalent to approximately 60–80% of the individual 1RM to allow for completion of 10–15 repetitions (Nelson et al., 2007) per set. In recent years, these recommendations have been subject to controversial debate, since a number of studies comparing the effectiveness of RT as performed at different intensities of load have suggested that, in both young and senior adults, training at lower intensities may be similarly if not equally effective in invoking protein synthesis and associated gains in muscle size and

**Fig. 2.** Forest plot (a) reflecting effect sizes and 95% confidence intervals from the 15 studies reporting training-induced changes in strength. Funnel plot (b) relating the studies’ effect sizes to the inverse of their standard error. (a) Effect sizes for non-matched and work-matched studies are shown separately. The values over the markers reflect the statistical weight assigned to the study. Note the estimated total population effect at the bottom in the dark grey row. (b) Note that one study (Singh et al., 2005) disproportionally shifts the plot to the right, providing evidence of asymmetry (Egger’s test: $t = 2.294$, $P = 0.041$). Favors LIT, favors light-intensity resistance training; Favors HIT, favors high-intensity resistance training.
strength (Pruitt et al., 1995; Taaffe et al., 1996; Bemben et al., 2000; Burd et al., 2010; Mitchell et al., 2012). One explanation for the discrepant findings reported in studies comparing different RT regimen lies in the fact that differences in the total mechanical work performed or the degree of fatigue induced by the training interventions were often not accounted for. Such differences may bias the comparison of the effectiveness of training at unequal intensities of load. This notion is supported by recent evidence to show that gains in muscle mass and strength following training with either heavier or lighter loads were similar, when training protocols were matched for mechanical work (Léger et al., 2006; Alegre et al., 2015). When assisted by artificial blood flow restriction to invoke early muscular fatigue, even intensities as low as 15–20% of 1RM may suffice to provoke substantial training effects (Loenneke et al., 2012). Data obtained in young and previously untrained individuals...
therefore suggest that training with lower loads may represent a suitable alternative to conventional, heavy-load RT (Schoenfeld, 2013a).

Elderly cohorts precluded from conventional RT because of strict contraindications for high-intensity RT, such as uncontrolled hypertension or cardiovascular diseases (Williams et al., 2007; Thompson et al., 2013), would benefit the most from the possibility to lower training loads. Moreover, lower load RT might represent an interesting alternative for those suffering from degenerative joint diseases, such as advanced osteoarthritis, since high stresses imposed on joint structures and connective tissues may cause pain and potentially further damage to joints (Segal et al., 2014). Research further suggests that training at lower intensity may generally be safer as the risk of training-related injuries has been found to increase with higher loads even in healthy, experienced weightlifters (Schoenfeld et al., 2014a).

Ultimately, elderly subjects have reported to prefer training with lighter loads (King et al., 1991), which may explain why training adherence increases when intensities of load are reduced (Perri et al., 2002).

Across all studies included, our results demonstrate that RT may provoke substantial increases in muscle strength even if performed at lower than conventionally recommended intensities of load (mean effect size of \(-0.66\) favoring training at low loads over no training). These findings are in agreement with the conclusions drawn in another recent meta-analysis including young and healthy individuals (Schoenfeld et al., 2014b) and appear to contest the concept of a “strength-endurance continuum” that implicates that improvements in the ability to exert maximal forces require training at high (near maximal) loads (Campos et al., 2002). When comparing the effectiveness of different training intensities, however, our data further suggest that strength gains are even larger when training with loads reaching or exceeding 80% of 1RM (total population effect size 0.43 favoring conventional RT). It is noteworthy that this analysis included data by Singh et al. (2005) that resulted in an effect size \(r = 0.990\) that was substantially larger than those calculated for all other studies included, and consequently appeared as an outlier on the Funnel plot presented in Fig. 2(a). The reason why high-intensity training was favored substantially more strongly in this study may lie in the particularly low loads (20% 1RM) chosen for the low-intensity training group; such training stimuli are likely to be insufficient to provoke strength gains. In addition, this investigation was carried out in subjects suffering from clinical depression, whose response to RT may differ from that typically observed in healthy elderly cohorts. This observation notwithstanding, our results still lend support to traditional training recommendations (ACSM, 2009a; Peterson & Gordon, 2011) and confirm previously drawn conclusions that high-intensity RT is a more potent stimulus to induce strength gains as compared with training at lower intensities (Peterson et al., 2010; Steib et al., 2010; Raymond et al., 2013). Expanding on these earlier works, we made a particular effort to consider bias that might result from unequal overall training loads by running independent analyses only on those studies where the overall amount of mechanical work performed \((sets \times repetitions \times load)\) was matched between training groups. Effect sizes reflecting the differences between strength gains were smaller by \(-37\%\) in work-matched studies although pooled data still favor conventional over lighter load RT \((population effect size \sim 0.30)\). As a word of caution, however, it should be noted that the results reported in three papers by Fatouros et al. (2005a, b, 2006) yielded composite effect sizes that were significantly larger \((r \sim 0.67)\) than those calculated for all other work-matched studies. After exclusion of these studies, the population effect size decreased to \(\mu = 0.15\) in favor of RT at higher intensities of load. These results suggest that the deficits in terms of mechanical stimuli that result from usage of lower intensities of load may, to a large extent, be compensated by increasing the repetition number and, thus, the training-associated metabolic stress (Schoenfeld, 2013b).

To the best of our knowledge, this review is also the first to investigate the role of training intensity in invoking muscle hypertrophy in senior cohorts. The results of seven studies to report changes in muscle mass, volume, cross-sectional area or circumference indicate that, at advanced age \((mean \sim 68 years)\), both conventional and lower load RT are largely ineffective in inducing total muscle growth. In the light of previous studies reporting substantial hypertrophy at the single muscle fiber level \([10-62\% \ after \ 9-52 \ weeks \ of \ training \ (Hunter \ et \ al., \ 2004)]\), these results may seem surprising. It should be noted, however, that the capability of senior subjects’ muscles to grow at the whole muscle level may be impaired because of age-related motor unit loss and a concomitant accumulation of intramuscular non-contractile tissues (Csapo et al., 2014). In agreement with this notion, previous studies in elderly cohorts found only very modest \(<10\%)\) training-associated gains in quadriiceps femoris volume or cross-sectional area (Hakkinen et al., 1998; Reeves et al., 2004; Verdijk et al., 2009; Stewart et al., 2014).

Relating the results of the present meta-analysis to the decreases of muscle strength and mass reported in the literature, the decline of strength is about 10% from the fifth decade onwards (Vandervoort, 2002). Changes in local muscle mass show decreases of about 15% in mid-thigh cross-sectional area from the age of 65 to 77 years (Frontera et al., 2000). The average strength increases in the “work-matched” studies of our review were 43% and 35% for the high and light-moderate loads, respectively. This implies that, at older age (mean age \sim 68\ years), resistance training partly reverses the age-associated loss of muscle function, returning strength to the levels of subjects 20 years younger.
Regarding the efficacy of resistance training to restore muscle mass, our data showed that the increases in appendicular muscle mass reported in studies classified as "work-matched" were 11% and 9% (high and lighter loads, respectively). Thus, it seems that, while the hypertrophic potential seems generally reduced in senior muscles, resistance training as started at advance age may still attenuate although not completely reverse age-associated loss in muscle mass (Frontera et al., 2000).

**Perspectives**

The present synopsis of current literature demonstrates that RT at lower than traditionally recommended intensities of load (~45% 1RM) may suffice to induce substantial gains in muscle strength in elderly cohorts. Training with heavier loads may still be required to maximize strength gains, although the analysis of a subset of studies in which training was matched for mechanical work suggests that greater training volumes may largely compensate for lower intensities. Both RT at high (~80% 1RM) and lower intensities of load provoke only minor increases in total muscle size, which indicates that the hypertrophic potential of skeletal muscles is blunted at older age. The present results could be useful for the application of RT programs in elderly people precluded from training with high intensities.

**Key words:** Low-load exercise, muscular adaptations, sarcopenia, resistance exercise, hypertrophy.

**Acknowledgements**

This work was supported in part by the Spanish Department of Health and Institute Carlos III (grant number RD12/043/0002) of the Government of Spain [Spanish Net on Aging and Frailty; (RETICEF)]. The authors wish to thank Dr. Alejandro Santos-Lozano for his useful statistical advice.

**References**


Morse CI, Thom JM, Mian OS, Birch KM, Narici MV. Gastrocnemius specific force is increased in elderly males following a 12-month physical training programme. Eur J Appl Physiol 2007: 100: 563–570.


Schuenke MD, Herman J, Staron RS. Preponderance of evidence proves “big” weights optimize hypertrophic.