Resistance Exercise for Muscular Strength in Older Adults: A Meta-Analysis

Mark D. Peterson, Ph.D., Matthew R. Rhea, Ph.D., Ananda Sen, Ph.D., and Paul M. Gordon, Ph.D., MPH

1Department of Physical Medicine and Rehabilitation, University of Michigan
2Department of Interdisciplinary Studies, AT Still University/Arizona School of Health Sciences
3Department of Statistics, University of Michigan

Abstract

Purpose—The effectiveness of resistance exercise for strength improvement among aging persons is inconsistent across investigations, and there is a lack of research synthesis for multiple strength outcomes.

Methods—The systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) recommendations. A meta-analysis was conducted to determine the effect of resistance exercise (RE) for multiple strength outcomes in aging adults. Randomized controlled trials and randomized or non-randomized studies among adults ≥50 years, were included. Data were pooled using random effect models. Outcomes for 4 common strength tests were analyzed for main effects. Heterogeneity between studies was assessed using the Cochran Q and I² statistics, and publication bias was evaluated through physical inspection of funnel plots as well as formal rank-correlation statistics. A linear mixed model regression was incorporated to examine differences between outcomes, as well as potential study-level predictor variables.

Results—Forty-seven studies were included, representing 1079 participants. A positive effect for each of the strength outcomes was determined however there was heterogeneity between studies. Regression revealed that higher intensity training was associated with greater improvement. Strength increases ranged from 9.8 – 31.6 kg, and percent changes were 29 ± 2, 24 ± 2, 33 ± 3, and 25 ± 2, respectively for leg press, chest press, knee extension, and lat pull.

Conclusions—RE is effective for improving strength among older adults, particularly with higher intensity training. Findings therefore suggest that RE may be considered a viable strategy to prevent generalized muscular weakness associated with ageing.

Keywords
Sarcopenia; Resistance Training; Aging; Systematic Review
Introduction

Muscular weakness plays a principal role in the pathogenesis of frailty and functional impairment that occurs with aging, and contributes to numerous disease processes. Maximal strength capacity reaches a peak sometime around the second or third decade of life, and by the fifth decade, begins a gradual decline (Larsson, et al., 1979; Lindle, et al., 1997; Metter, et al., 1997; Narici, et al., 1991; Vandervoort, et al., 1986). This deterioration, which is typically attributed to diminished levels of activity or disuse/immobilization due to disease, has been documented primarily through cross-sectional research, and appears to increase in severity after the age of 65 (Baumgartner, et al., 1998). Although losses of strength are rarely tracked longitudinally (Aniansson, et al., 1986; Bassey, et al., 1993; Frontera, et al., 2000; Kallman, et al., 1990), existing epidemiological studies report a significantly higher prevalence across each decade of late adult life. Sarcopenia and muscular weakness are not considered to be “disease” states, but rather conditions which translate to acute functional deficit and disability, as well as related comorbidity and mortality (Ruiz, et al., 2008). Moreover, increased longevity has led to a higher frequency of sarcopenia, and respective escalating health care expenditures for complications associated with declines in functional health and loss of independence (96).

Numerous investigations have identified a disparate decline of strength and muscle mass, indicating that these age-related debilities are to some extent, independent (Klitgaard, et al., 1990; Lynch, et al., 1999; Young, et al., 1985). However, since strength and muscle mass do not decrease concurrently, strength may be a superior indicator of muscular dysfunction (Doherty, 2003; Klein, et al., 2001). Indeed, longitudinal data suggest that muscle strength is a robust predictor of functional decline that may occur during aging (Pendergast, et al., 1993; Rantanen, et al., 1999; Rantanen, et al., 1999), and is an important physiological attribute for maintenance of mobility and movement efficiency. Since strength capacity appears to also be indicative of disability (Janssen, et al., 2002; Visser, et al., 2002), resistance exercise (RE) may serve as an effective mode of exercise to directly improve functional capacity. There is strong evidence to suggest that muscle weakness is a treatable cause of disability, and that aging persons with early-onset deterioration are probably the most likely to benefit from strategic interventions (Evans, et al., 1996; Frontera, et al., 1988; Hakkinen, et al., 1995; Hurley, et al., 1995). Specifically, RE is considered to be a safe and effective method for increasing strength and lean muscle tissue in young (Hubal, et al., 2005; Lowndes, et al.) and older adults (Fiatarone, et al., 1990; Frontera, et al., 1988; Hakkinen, et al., 1998; Hakkinen, et al., 2001; Reeves, et al., 2004; Vincent, et al., 2002; Welle, et al., 1995). Even after short bouts of resistance training protocols, aging subjects may experience improvements in protein synthesis rate and neuromuscular adaptation that are comparable to that of younger cohorts, despite a much lower pre-exercise rate (Holviala, et al., 2006; Newton, et al., 2002; Roth, et al., 2001; Yarasheski, et al., 1993). These findings imply that disuse may actually be the underlying reason for muscle atrophy and weakness, rather than aging, per se.

The effectiveness of RE for strength improvement among aging persons is inconsistent across investigations. While much research has examined strength increases accompanying single-cohort interventions, most have examined only one or two training programs, providing only a glimpse of the overall dose-response relationship. Debate concerning the appropriateness of RE among older individuals has been cultivated by questions of the general efficacy and safety for this population. There are very few published accounts that examine the overall benefit of RE for strength in aging persons while considering a continuum of dosage schemes, treatment durations, and/or age ranges on longitudinal strength adaptation. As a result, it is difficult to evaluate the treatment effects coinciding with these factors. Further, although the existing body of evidence regarding the utility of resistance exercise for strength improvements among older adults has recently been deemed to be supported by the highest category of evidence (i.e. “Evidence Category A.”) (Chodzko-Zajko, et al., 2009), a systematic review to assess treatment
effects across multiple strength measures, and potential moderating variables more
generalizable to RE prescription, is yet to be completed. To date, the most comprehensive
reviews related to this topic have either limited the analysis of strength to a single measure (i.e.
knee extension) (Latham, et al., 2003, Latham, et al., 2004, Liu, et al., 2009), or have compared
multimodal exercise regimens for general changes in overall functional capacity (Baker, et
al., 2007), not specific to strength outcomes. Therefore the purposes of this review and meta-
analyses were to examine the effects of resistance exercise among older adults for multiple
upper- and lower-body strength outcomes, and across multiple dosing schemes.

Methods

This meta-analysis was conducted in accordance with the recommendations and criteria as
outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses
(PRISMA) statement.

Criteria for study consideration: Types of studies and participants

All resistance exercise interventions that included the primary outcome related to muscular
strength were included in the original article acquisition. Any randomized controlled trials
(RCTs) or quasi-randomized clinical trials meeting the subsequent specifications were
included. Randomized or non-randomized treatment studies (non-RCTs) that examined
intervention treatments using stratified young versus older participants, or aged men versus
women were also eligible for inclusion in the analyses. The type of participants included older
men and women, residents in institutions or hospitals, or senior-citizen communities. Trials
were included if the mean age of participants was ≥50 years, but excluded if participants aged
<50 years were enrolled. The physical health of participants ranged from fit and healthy to
frail or disabled older people, and/or people with identified diseases or health problems. Similar
to previous reviews (Latham, et al., 2003, Latham, et al., 2004), inclusion of participants with
a range in age and health complications was critical to increase external validity and
generalizability of results.

Criteria for study consideration: Types of interventions and outcome measures

Trials that had at least one group of participants who received RE as a treatment were
considered for inclusion. Training could have taken place in group exercise programs (e.g. in
commercial health clubs, etc.), individual personal training arrangements, and/or senior
community-based settings. For these analyses, RE was defined as a strength training regimen
that included specific exercises for the whole body. Training protocols were classified in
accordance with the American College of Sport Medicine Position Stand on Progression
Models in Resistance Training for Healthy Adults (Kraemer, et al., 2002). Study inclusion for
functional improvements was limited to four discrete measurements of maximal strength
capacity, including one-repetition maximum (1RM) tests for machine leg press, chest press,
knee extension, and lat pull (i.e. lat pulldown and/or lat row). These tests were selected due to
documented reliability for assessments, as well as reported prevalence in the literature. Other
methods of assessing muscular fitness (e.g. power, endurance, etc.), functional deficit or
performance were not included in the analysis.

Search strategy

Computerised searches of MEDLINE, EMBASE, PubMed, Web of Science, SPORTDiscus™,
Evidence Based Medicine Reviews Multifile (EBMR) databases, and Digital Dissertations
(accessed May, June, and July 2009) from their inception to July, 2009 were undertaken. Hand-
searching of key journals, reference lists and other sources was also undertaken. Studies
published in foreign language journals were not included. Abstracts and citations from annual
scientific conferences relating to exercise science or gerontology were not examined due to

Ageing Res Rev. Author manuscript; available in PMC 2011 July 1.
paucity of requisite data. The preliminary search yielded over 5,000 relevant abstracts and citations. Full texts of over 400 articles were obtained and examined by the primary reviewer (MP).

**Study eligibility and data extraction**

A specific coding tool was developed from previous quantitative reviews, as well as suggestions from meta-analysis experts, to record information pertaining to the study source, participants, experimental characteristics, and outcomes. Although all eligible studies in this investigation shared a common directive to examine the effectiveness of RE for older adults, several studies examined different hypotheses (e.g. younger versus older (Roth, et al., 2001)). Only data from the older participants were coded for analysis. For all interventions included in this meta-analysis, each treatment group was considered a one-group pretest-posttest intervention. Therefore coding of strength change was only carried out for groups receiving the RE treatment, and all data reported from controls were disregarded. If data were not adequately reported for the purpose of quantitative pooling, the author of correspondence was contacted for additional information. If the authors could not be contacted or if the information was no longer available, the trial was excluded from analyses.

**Assessment of reviewer agreement and risk of bias for included studies**

Inter-reviewer disagreements were resolved by consensus. The agreement rate prior to amending any such discrepancies was assessed using the kappa statistic (Donner, et al., 1996), and determined to be 0.93. In the case of inadequate information contained in the manuscript, the lead reviewer (MP) sought clarification from study authors. To assess for evidence of publication bias, Begg's funnel plots were examined (Begg, et al., 1989). This process included the visual inspection of scatterplots for each measure against its standard error. This process is necessary to account for the “file drawer problem,” which is the potential effect that published studies are inherently biased due to a greater likelihood of significant results, whereas non-significant results are more likely to not be submitted for publication and be placed in a file drawer. The use of published studies could represent an inflated account of the effect between treatment and outcome, and thus it is necessary to test for bias. As a subsequent check of publication bias, the test of Egger et al. (Egger, et al., 1997) was incorporated. This linear regression method quantifies the bias captured by the funnel plot, and more specifically, the standardized effect is regressed on precision (i.e. inverse of standard error) (Borenstein, et al., 2009). This is a formal statistic that is intended to assess the same assumption as the funnel plot, and may be used as a “cross-check” to the physical inspection of the data.

**Tests for Heterogeneity**

Heterogeneity between studies was assessed using the Cochran $Q$ statistic (Cochran, 1954). This test is appropriate in larger meta-analyses and uses the sum of squared deviations of the study-specific estimates derived from the pooled estimate, and weights the contribution of each study. $P$ values were obtained by comparing the $Q$ statistic with a $\chi^2$ distribution and $k$-1 degrees of freedom, in which $k$ represents the number of studies included. Heterogeneity refers to the existence of variation between studies, on the main effects being evaluated. Since heterogeneity is to a certain extent, inevitable in meta-analytic research, there is ample debate regarding the utility of assigning statistical significance to this computation. Thus we also incorporated the $I^2$ statistic, using the following equation:

$$I^2 = \left(\frac{Q}{df}/Q\right) \times 100\%$$
With this method $I^2$ ranges from 0% - 100% and values greater than 50% are used to indicate meaningful heterogeneity.

**Data Syntheses**

Treatment effects for muscular strength capacity were calculated for each study following the extraction/coding of change scores and standard deviations. Specifically, the standard deviation (SD) of change was needed to calculate the effect size, and for many of the studies this value was not reported. Rather, the majority of studies obtained for this analysis included the SDs for the baseline and postintervention strength outcomes, or in many cases the standard errors of the mean. In the event that the study reported exact $P$ values for the change in strength outcome, the SD of change was computed. However, for those studies which did not report exact $P$ values, the SD of change was calculated using the baseline and post-intervention SDs, as well as the within-participant bivariate correlation of strength measures using the following equation:

$$SD\ Change = \sqrt{[(SD_{pre})^2 + (SD_{post})^2] - 2 \times corr(pre, post) \times SD_{pre} \times SD_{post}}$$

For every article included, authors were contacted in an effort to retrieve raw data for the calculation of the within-participant baseline and postintervention strength correlations, or the specific and respective $r$ (correlation) values. For those studies in which authors could not be reached, $r$ was imputed using the mean of the correlations available for 15 of the included articles. This resulted in using a within-participant correlation of $r = 0.95$ for the ones that were missing. This allowed for the computation of effect sizes for all cohorts included in the metaanalysis, as recommended by Follmann et al (Follmann, et al., 1992).

**Effect Sizes**

The analysis of pooled data was conducted with a fixed and random-effects model. A fixed effects model is used with the assumption that sampling error is random, and is the primary cause of variation in the summary effect. Conversely, a random effects model is incorporated when the underlying assumption is that the effect across studies is randomly situated about a central value (Borenstein, et al., 2009). For each of the four strength measures, forest plots were generated to illustrate the study-specific effect sizes, and the respective 95% confidence intervals (CI). Combining estimates then allowed for the assessment of a pooled effect, as has been previously described (Richardson, et al., 2008), in which the reciprocal of the sum of two variances were accounted for, including: (1) the estimated variance associated with the study, and (2) the estimated component of variance due to variation between studies. In each study, the effect size for the intervention was calculated by the difference between the means of the post-test and pre-test at the end of the intervention. The study specific weights were derived as the inverse of the square of the respective standard errors.

In order to compare the pre-post change between types of strength outcomes, the data were analyzed under a linear mixed models framework, in which the outcomes were calculated as mean percent change in strength. The study-level independent variables used in the regression model were strength measure type, indicator for study design, gender, age-group, length of training, average intensity (4 subgroups), and average training volume (4 subgroups). An interaction between strength measure types was also included in the model to assess any differential age effect across measure types. A random study effect accounted for the clustering within study. Further, each study was weighted by the inverse of the square of the study-specific standard error corresponding to the outcome. The standard deviation (standard error multiplied by $\sqrt{\text{samplesize}}$) for percentage error was approximated by the following equation:
Specifically, CV denotes the coefficient of variation equaling the ratio of standard deviation and mean at the respective time points, and R is the ratio of the post and pre strength measures. For post-hoc comparisons of strength measures, Bonferroni adjustments were used to protect against an inflated Type 1 error. All statistical analyses were performed using STATA 10.0 (StataCorp LP, College Station, Texas), MINITAB 14.0 (Minitab Inc, State College, Pennsylvania), and SAS 9.2 (SAS Institute Inc., Cary, NC).

Results

The flow of article search and selection, from “potentially relevant” to final inclusion is depicted in Figure 1.

Study characteristics


Subject characteristics


\[ \text{SD PerChange}=R \times \sqrt{[(CV_{\text{pre}})^2+(CV_{\text{post}})^2 - 2 \times \text{corr}(\text{pre}, \text{post}) \times CV_{\text{pre}} \times CV_{\text{post}}]} \]

Treatment characteristics

Length of training ranged from 6 to 52 weeks (mean duration = 17.6 ± 8.6 weeks), frequency from 1 to 3 times per week (mean = 2.7 ± 0.5 days/week), and intensity from 40% to 85% of 1 repetition maximum (mean = 70 % ± 12.7 1RM). The number of sets per exercise session ranged from 1-6 sets for each individual muscle (mean = 2.5 ± 1.0 sets), while the number of exercises performed ranged from 5-16 (mean = 8.3 ± 2.1 resistance exercises). The within-group number of repetitions performed for each set ranged between 2 and 20 (mean = 10 ± 2.6 repetitions), while the rest period between sets ranged from 60 to 360 seconds (mean = 110 ± 25 seconds). Compliance, defined as the percentage of exercise sessions attended, ranged from 85 to 100%.

Main Effects

Each strength measure was independently assessed through meta-analytic procedure, and is presented sequentially. Many trials reported more than a single outcome (Table 1). Due to the significant heterogeneity of the data based on Cochran's $Q$ and $I^2$, a random-effects model was incorporated for each type of strength measure.


When analyzed separately, sub-group analyses for the main effect revealed non-significant differences between randomized-controlled trials (RCTs) (pooled estimate = 34.38 kg; 95% CI, 27.74 kg to 41.02 kg) and randomized or non-randomized treatment studies (non-RCTs) (pooled estimate = 28.99 kg; 95% CI, 23.47 kg to 34.52 kg), despite significant increases in strength from pre- to post-intervention, for cohorts in both designs ($p < 0.001$).


Sub-group analyses revealed non-significant differences between RCTs (pooled estimate = 9.97 kg; 95% CI, 7.57 kg to 12.38 kg) and non-RCTs (pooled estimate = 9.68 kg; 95% CI, 7.79 kg to 11.58 kg), despite significant increases in strength from pre- to post-intervention, for cohorts in both designs ($p < 0.001$).


Sub-group analyses revealed non-significant differences between RCTs (pooled estimate = 12.73 kg; 95% CI, 10.38 kg to 15.08 kg) and non-RCTs (pooled estimate = 9.83 kg; 95% CI, 7.93 kg to 11.73 kg), despite significant increases in strength from pre- to post-intervention, for cohorts in both designs ($p < 0.001$).


Sub-group analyses revealed non-significant differences between RCTs (pooled estimate = 10.97 kg; 95% CI, 8.21 kg to 13.73 kg) and non-RCTs (pooled estimate = 9.7 kg; 95% CI, 7.26 kg to 12.14 kg), despite significant increases in strength from pre- to post-intervention, for cohorts in both designs ($p < 0.001$).

**Publication Bias**

Examination of the Beggs funnel plots for leg press, chest press, and lat pull demonstrated considerable symmetry, suggesting that there was no significant publication bias. Results from the Egger’s test further confirmed no evidence of publication bias ($p = 0.12 – 0.74$), for each of these main outcomes. Conversely, for knee extension, inspection of the funnel plots (Figure 6) as well as results from the Egger’s test ($p < 0.001$) suggested positive evidence of publication bias. In the plot, the pair representing actual strength change and its standard error for each study is represented by circles. A horizontal line at the height equaling the variance-weighted, meta-analytic effect estimate is drawn with the two walls of the cone representing the confidence interval limits (estimate +/- 1.96* standard error), about each study-specific effect.
Asymmetry on the right of the graph (i.e. where studies with high standard error are plotted) may indicate evidence of publication bias.

**Difference across measures**

Significant differences were obtained in percent changes of strength across measure types. The mean ± standard error of percent changes in strength were 29 ± 2, 24 ± 2, 33 ± 3, and 25 ± 2 respectively, for leg press, chest press, knee extension, and lat pull. In post-hoc comparisons, the percent change in knee extension turned out to be significantly greater than either chest press (Bonferroni adjusted p-value = 0.003) or lat pull (Bonferroni adjusted p-value = 0.007). The only between-study predictor that had a significant association with percent strength change was training intensity. The average percent change in strength for an incremental increase in intensity subgroup was found to be 5.3% (standard error = 0.9, p < 0.001).

**Discussion**

The primary results of this investigation suggest a robust, significant association between resistance exercise and upper and lower body strength improvement among older individuals. From a public health perspective, these findings confirm the value of full-body RE for the prevention or treatment of age-related declines in muscle function, which may in turn serve as a safeguard against disabulment. In particular, we observed significant main effects for lower-body (i.e. leg press = 31.63 kg (29%); knee extension = 12.08 kg (33%)) and upper-body (i.e. chest press = 9.83 kg (24%); lat pull = 10.63 kg (25%)) strength capacity, following RE interventions. These findings bear clinical significance, considering the exaggerated strength decline that occurs among sedentary individuals after the age of 50 years (Borges, et al., 1989, Larsson, et al., 1979, Lindle, et al., 1997, Metter, et al., 1997, Narici, et al., 1991, Vandervoort and McComas, 1986), as well as the subsequent contribution of strength deficit to disability and movement impairment (Pendergast, et al., 1993, Rantanen, et al., 1999, Rantanen, et al., 1999).

Previous meta-analyses conducted on this population have restricted the examination of muscle function following “progressive resistance training (PRT)” to changes in knee extensor strength capacity, for the purpose of reducing the risk of “clinical heterogeneity” (Latham, et al., 2003, Latham, et al., 2004, Liu and Latham, 2009). This criterion is rational, given the high frequency with which knee extension strength is reported in the literature, as well as the relevance of lower-limb strength to locomotion, activities of daily living, and risk of slip-and-fall accidents (Pijnappels, et al., 2008). Moreover, due to the large cross-sectional area of the quadriceps, the degree of strength decline and muscle atrophy during aging appears to exceed that of the upper-body (Doherty, 2003), and thus an intervention model for positive effects on lower-limb strength provides salient inference for enhancement of overall functional capacity. However, the aforementioned analyses yielded significant statistical heterogeneity in the strength data that could not be attributed to differences in the study quality, participant characteristics or the exercise program variables (Latham, et al., 2003, Latham, et al., 2004).

In order to improve external validity and clinical generalizability associated with prescription of RE in older adults, the current analysis separated and examined the four most frequently tested strength outcomes (i.e. leg press, chest press, knee extension, and lat pull), and subjected each measure to individual meta-analytic synthesis and post-hoc analysis. These measures were chosen due to the high frequency of reporting in the literature, and ultimately because the aggregate represents a superior indicator of whole body functionality. To our knowledge, this is the first meta-analysis to synthesize data from full-body resistance training programs conducted on older men and women, and to report the main effects from multiple strength outcomes.
Similar to previous reviews (Latham, et al., 2003, Latham, et al., 2004, Liu and Latham, 2009), we found significant heterogeneity in the data for each outcome estimate, and thus a random effects model was necessary to analyze pooled effect sizes. Based on the linear mixed model regression and sub-group analyses, this heterogeneity could not be explained by differences in age, indicator for study design, length of training, or resistance training volume. Although there was variability in the individual study characteristics, nearly every resistance exercise protocol conformed to the American College of Sports Medicine (ACSM) recommendations for resistance exercise in older adults (ACSM, 1998). Further, according to the mixed model regression, strength improvement across all four outcomes shared a significant, positive relationship with resistance training intensity, suggesting that higher intensity RE programs are superior for strength improvement. In particular, training intensity ranged from approximately 40-90% of 1RM. Based on the a priori designation of low intensity (< 60% 1RM), low/moderate intensity (60-69% 1RM), moderate/high intensity (70-79% 1RM), and high intensity (≥ 80% 1RM), the mean change in relative strength (i.e. percent from pre to postintervention) for an incremental increase in intensity subgroup was nearly 5.5%. This is similar to previous reviews (Latham, et al., 2003, Latham, et al., 2004) in which subgroup analysis demonstrated that higher intensity training was associated with greater strength improvement among older populations, as compared to low and moderate intensity training.

The results from the linear mixed model on RE prescription variables did not identify any significant relationships between intervention duration or training volume, and subsequent strength effects. This is contrary to data from independent studies which have demonstrated the superiority of higher dosage RE for strength adaptation, longitudinally (Galvao and Taaffe, 2005). For the current analysis it is conceivable that the overall lack of substantial variability in training regimens across program models (Table 1) may have convoluted these results. Specifically, without a sufficient number of cohorts to stratify into trichotomous categories for each of the variables (e.g. high-, medium-, and low-volume), it is not feasible to effectively pool data to estimate individual main effects.

Current published RE recommendations are considerably different for young and middle-aged healthy adults (ACSM, 2009, Kraemer, et al., 2002), as compared to those for elderly populations (ACSM, 1998, Chodzko-Zajko, et al., 2009, Nelson, et al., 2007). The majority of published intervention studies and subsequent recommendations for younger adults have incorporated suggestions of periodization to promote enhanced adaptation of strength, whereas no such recommendations have been endorsed for the aging population. Since most of the studies in the current analysis followed the basic recommendations by the ACSM (ACSM, 1998, ACSM, 1998), the only requisite “progression” in training prescription was that of the absolute training load (i.e. “Progressive Resistance Training (PRT)”). However, merely increasing training load over time may not be sufficient beyond a certain point, as this represents an inevitable reliance on the same relative intensity (i.e. per maximal strength capacity). It is therefore conceivable that current findings misrepresent the true relationship between training volume dosages, and respective strength adaptation-potential of aging men and women. Although results from the current study confirm that resistance exercise is very effective for increasing strength capacity in this population, further research is warranted that examines interventions with hierarchical resistance exercise prescription models, and variable volume dosage schemes among healthy older adults.

Previous comparisons between young and/or middle-aged men and women and healthy elderly samples have yielded conflicting results (Hakkinen, et al., 1998, Holviala, et al., 2006, Lemmer, et al., 2000). Several investigations have demonstrated similar strength improvements (Hakkinen, et al., 1998, Holviala, et al., 2006, Newton, et al., 2002), whereas others suggest greater strength adaptation among younger cohorts (Lemmer, et al., 2000, Macaluso, et al.,
For the current review, regression analyses failed to identify an association between gender or age and strength main effects, and thus imply that the potential adaptive-response is significant for both men and women, across each decade of late adult life. Although significant strength adaptation is possible in the “oldest old” (Kryger, et al., 2007), it may be expected that the benefits of early strength acquisition will translate to superior preservation of functional movement capacity and instrumental activities of daily living, prevention of disability, and maintenance of independence and autonomy. In fact, as has recently been suggested, a reasonable degree of variation in muscular atrophy and weakness between older individuals may in part be attributable to the peak attained earlier in life (Sayer, et al., 2008). However, as a cautionary statement it should be noted that the vast majority of these current data were derived from healthy older adults, < 80 years old. Certainly more research is needed to ascertain the influence of RE for specific disease outcomes, and across a broader spectrum of age categories.

Strict inclusion criteria related to study design for meta-analysis has been widely debated. Systematic reviews generally employ only randomized controlled trials due to the removal of increased risk of bias in nonrandomized samples. However, there have been several recent reviews that have demonstrated no differences in effect sizes between randomized and nonrandomized trials (Benson, et al., 2000, Concato, et al., 2000), and there is ample debate regarding the value of this quality index for meta-analysis study inclusion (Balk, et al., 2002). With regard to assessing indicators of study “quality,” no acceptable scale currently exists for examining the quality of exercise intervention research. Moreover, the Cochrane Collaboration has recently updated guidelines for systematic reviews, and has recommended against the use of quality scales due to an overall lack supporting evidence and validity (Higgins, et al., 2009). Therefore, sub-analysis for overall study quality was not carried out for the current investigation, although previous reviews have reported the overall lack of quality among the majority of resistance training literature for older adults (Latham, et al., 2004).

Although meta-analyses usually solicit articles from large databases there is typically a limitation with regard to research design quality. In the present study, only 25 out of 47 studies incorporated randomized controlled designs (RCTs). The remaining studies were classed as non-RCTs that often did not have a control group or involved placing subjects into groups by convenience, availability, or age/gender stratification. Sub-group analyses demonstrated no significant differences for the main effect between RCTs and non-RCTs, and in all cases, the main effects for non-RCTs were actually slightly less than that demonstrated among RCT designs. Certainly, the body of knowledge pertaining to the effects of RE on muscular strength among older adults would be enhanced if there were more investigations with a true experimental design. Moreover the acceptance of RE as a viable preventive or treatment option should elicit larger available samples from which to recruit, and thus reduce the risk of low-powered studies.

As with all meta-analysis, general limitations include the comparison of aggregate outcomes that do not necessarily assess the exact same construct (i.e. “comparing apples and oranges”), the process of search and retrieval for eligible articles, and the potential influence of publication bias (Rosenthal, 1991). For this study, every attempt was made to examine outcomes of similar construct. However, whereas previous reviews have restricted the analyses of strength effects to knee extension (Latham, et al., 2003, Latham, et al., 2004, Liu and Latham, 2009), we pooled data across various strength measures, in an effort to enhance external validity. Cooper et al. (Cooper, et al., 1994), also reported a “retrieval bias” in which criteria for soliciting articles is potentially inadequate. For example, in this analysis, retrieval bias may have been an issue due to the selection and retrieval of articles published only in the English language. Despite that over 5000 abstracts were reviewed, the selection of articles for inclusion could have potentially resulted in more publications if this criterion was not established. Publication bias has also...
been a longstanding concern with meta-analysis and systematic reviews. In essence, the use of published data may represent an inflated account of the effect between treatment and outcome. To this end, many reviewers have evaluated publication bias through the visual examination of funnel plots. Although a common practice, the use of funnel plots has received some degree of criticism (Tang, et al., 2000). At present there is not a uniform consensus on how these plots should be constructed for examining exercise intervention literature, and thus asymmetrical funnel plots should be interpreted cautiously. For the current analyses, the only evidence of publication bias determined through these tests was for the main effects of knee extension strength.

Ultimately, this meta-analysis cannot infer a causal-effect for disability or functional deficit, per se. Since age-related muscular weakness is a collection of interrelated deteriorations that coincide with chronological aging, the improvement of strength capacity through RE is a viable preventive strategy to complement other medical and lifestyle interventions. Muscle weakness is associated with specific functional deficits, disease comorbidity, and escalating health care expenditures, and thus more studies are needed to directly examine treatment options for these significant consequences. The investigation of resistance exercise for aging populations is a relatively contemporary focus of research scholarship that bridges multiple fields of study, and one that has potential to significantly impact quality of life for elderly individuals, as well as for overall public health enhancement.

Conclusions

Resistance exercise is an effective modality for aging men and women, and may elicit significant improvements in muscular strength capacity. The current analysis is supportive of previous findings to suggest a positive association between intensity of resistance exercise and degree of strength improvement. As strength decline is highly related with subsequent functional deficit and disease comorbidity, it is conceivable that improvements in this parameter would help to maintain independence, health, and overall well-being. At present, approximately only 27% of the U.S. population engages in leisure time resistance exercise activity (CDC, 2009). Participation rates are dramatically less for individuals over the age of 50, and are likely to be as low as 10% of adults > 75 years (CDC, 2009). Evidence from the current investigation confirms the efficacy of higher intensity RE as a robust prevention or treatment strategy for age-related losses of muscle function, and thus, increased public health efforts to facilitate the provision of this health behavior are certainly warranted.

Acknowledgments

This research is supported by the NIH, NICHD, NCMRR Grant #5-T32-HD007422. The authors thank Dr. Brent Alvar and Dr. Pamela Swan for their support and feedback during the conception of this project. No financial disclosures were reported by the authors of this paper.

References


Bassey EJ, Harries UJ. Normal values for handgrip strength in 920 men and women aged over 65 years, and longitudinal changes over 4 years in 620 survivors. Clin Sci (Lond) 1993;84:331–337. [PubMed: 8384956]


Figure 1.
Flow of papers through the review process.
Figure 2.
Forest plot of effect sizes and 95% confidence intervals for all 51 cohorts (32 studies) representing leg press, based on the random effects meta-analysis results.
Figure 3.
Forest plot of effect sizes and 95% confidence intervals for all 55 cohorts (36 studies) representing chest press, based on the random effects meta-analysis results.
Figure 4.
Forest plot of effect sizes and 95% confidence intervals for all 43 cohorts (28 studies) representing knee extension, based on the random effects meta-analysis results.
Figure 5.
Forest plot of effect sizes and 95% confidence intervals for all 38 cohorts (19 studies) representing lat pull, based on the random effects meta-analysis results.
Figure 6.
Begg's funnel plot for knee extension, with 95% confidence limits. Positive evidence of publication bias was found.
## Table 1

<table>
<thead>
<tr>
<th>Study Details</th>
<th>Subjects</th>
<th>PRT Intervention Details</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Duration (weeks)</td>
<td>Intensity (%)</td>
</tr>
<tr>
<td>Author, year</td>
<td>Design</td>
<td>Age (yr)</td>
<td>Gender</td>
</tr>
<tr>
<td>Ades, 1996</td>
<td>RCT</td>
<td>6</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>70</td>
</tr>
<tr>
<td>Ades, 2003</td>
<td>RCT</td>
<td>19</td>
<td>73</td>
</tr>
<tr>
<td>Ades, 2005</td>
<td>RCT</td>
<td>21</td>
<td>73</td>
</tr>
<tr>
<td>Bailor, 1996</td>
<td>non-RCT</td>
<td>9</td>
<td>61</td>
</tr>
<tr>
<td>Baumans, 2005</td>
<td>non-RCT</td>
<td>31</td>
<td>69</td>
</tr>
<tr>
<td>Benben, 2000</td>
<td>RCT</td>
<td>10</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>52</td>
</tr>
<tr>
<td>Beniamini, 1999</td>
<td>RCT</td>
<td>18</td>
<td>59</td>
</tr>
<tr>
<td>Bender, 2005</td>
<td>RCT</td>
<td>23</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>27</td>
<td>83</td>
</tr>
<tr>
<td>Borto, 2007</td>
<td>non-RCT</td>
<td>11</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9</td>
<td>66</td>
</tr>
<tr>
<td>Brochu, 2002</td>
<td>RCT</td>
<td>13</td>
<td>70</td>
</tr>
<tr>
<td>Campbell, 1994</td>
<td>non-RCT</td>
<td>12</td>
<td>65</td>
</tr>
<tr>
<td>Candon, 2006</td>
<td>non-RCT</td>
<td>29</td>
<td>59–76</td>
</tr>
<tr>
<td>Castaneda, 2001</td>
<td>RCT</td>
<td>14</td>
<td>65</td>
</tr>
<tr>
<td>de Vos, 2005</td>
<td>RCT</td>
<td>28</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td></td>
<td>28</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td></td>
<td>28</td>
<td>69</td>
</tr>
<tr>
<td>Fatouros, 2003</td>
<td>RCT</td>
<td>14</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td>71</td>
</tr>
<tr>
<td>Fatouros, 2006</td>
<td>RCT</td>
<td>14</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td>71</td>
</tr>
<tr>
<td>Figarreta, 2003</td>
<td>RCT</td>
<td>20</td>
<td>55</td>
</tr>
<tr>
<td>Study Details</td>
<td>Subjects</td>
<td>PRT Intervention Details</td>
<td>Outcomes</td>
</tr>
<tr>
<td>--------------</td>
<td>----------</td>
<td>--------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Author, year</td>
<td>Design</td>
<td>Treatment (Size)</td>
<td>Age (yr)</td>
</tr>
<tr>
<td>Galvao, 2005</td>
<td>non-RCT</td>
<td>Combined (7M, 5F)</td>
<td>24</td>
</tr>
<tr>
<td>Grewe, 2001</td>
<td>RCT</td>
<td>Combined (4M, 4F)</td>
<td>12</td>
</tr>
<tr>
<td>Hartman, 2007</td>
<td>non-RCT</td>
<td>Combined (not reported)</td>
<td>29</td>
</tr>
<tr>
<td>Haar, 2002</td>
<td>non-RCT</td>
<td>Combined (7M, 5F)</td>
<td>10</td>
</tr>
<tr>
<td>Hayekowsky, 2001</td>
<td>RCT</td>
<td>Combined (9M, 14F)</td>
<td>10</td>
</tr>
<tr>
<td>Hayekowsky, 2002</td>
<td>RCT</td>
<td>Combined (11M, 11F)</td>
<td>10</td>
</tr>
<tr>
<td>Henwood, 2006</td>
<td>RCT</td>
<td>Combined (6M, 9F)</td>
<td>23</td>
</tr>
<tr>
<td>Holvia, 2006</td>
<td>non-RCT</td>
<td>Combined (7M, 5F)</td>
<td>22</td>
</tr>
<tr>
<td>Humphries, 2000</td>
<td>non-RCT</td>
<td>Combined (9M, 14F)</td>
<td>21</td>
</tr>
<tr>
<td>Heflick, 2002</td>
<td>non-RCT</td>
<td>Combined (7M, 5F)</td>
<td>14</td>
</tr>
<tr>
<td>Ibanez, 2005</td>
<td>non-RCT</td>
<td>Combined (7M, 5F)</td>
<td>9</td>
</tr>
<tr>
<td>Igwebuike, 2008</td>
<td>RCT</td>
<td>Combined (11M, 11F)</td>
<td>14</td>
</tr>
<tr>
<td>Izquierdo, 2001</td>
<td>non-RCT</td>
<td>Combined (6M, 9F)</td>
<td>11</td>
</tr>
<tr>
<td>Izquierdo, 2003</td>
<td>non-RCT</td>
<td>Combined (6M, 9F)</td>
<td>11</td>
</tr>
<tr>
<td>Kalapothakis, 2005</td>
<td>RCT</td>
<td>Combined (4M, 7F)</td>
<td>11</td>
</tr>
<tr>
<td>Kraemer, 1999</td>
<td>non-RCT</td>
<td>Combined (4M, 7F)</td>
<td>9</td>
</tr>
<tr>
<td>Lemmer, 2001</td>
<td>non-RCT</td>
<td>Combined (4M, 7F)</td>
<td>11</td>
</tr>
<tr>
<td>Lemmer, 2007</td>
<td>non-RCT</td>
<td>Combined (11M, 10F)</td>
<td>10</td>
</tr>
<tr>
<td>Misico, 2003</td>
<td>RCT</td>
<td>Combined (6M, 7F)</td>
<td>13</td>
</tr>
<tr>
<td>Study Details</td>
<td>Subjects</td>
<td>PRT Intervention Details</td>
<td>Outcomes</td>
</tr>
<tr>
<td>---------------</td>
<td>----------</td>
<td>--------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Author, year</td>
<td>Design</td>
<td>Treatment Size (n) Age (yr) Gender</td>
<td>Duration (weeks) Intensity (%1RM) Volume (per exercise) Frequency (per wk)</td>
</tr>
<tr>
<td>Panton, 2001</td>
<td>RCT</td>
<td>11 Combined (5M, 6F) 67</td>
<td>12 70 3 sets 3 sessions</td>
</tr>
<tr>
<td>Pu, 2001</td>
<td>RCT</td>
<td>9 Combined (4M, 5F) 77</td>
<td>10 80 3 sets 3 sessions</td>
</tr>
<tr>
<td>Reeves, 2003</td>
<td>RCT</td>
<td>9 Combined (4M, 5F) 74</td>
<td>14 75 2 sets 3 sessions</td>
</tr>
<tr>
<td>Reynolds, 2007</td>
<td>non-RCT</td>
<td>13 Combined (6M, 7F) 66</td>
<td>16 75 2 sets 3 sessions</td>
</tr>
<tr>
<td>Reeves, 2003</td>
<td>non-RCT</td>
<td>9 Combined (4M, 5F) 69 M</td>
<td>24 75 1 set 3 sessions</td>
</tr>
<tr>
<td>Sharman, 2001</td>
<td>RCT</td>
<td>7 Combined (2M, 6F) 66 M</td>
<td>24 75 6 sets 2 sessions</td>
</tr>
<tr>
<td>Stewart, 2005</td>
<td>RCT</td>
<td>51 Combined (6M, 22F) 63 M</td>
<td>24 50 2 sets 3 sessions</td>
</tr>
<tr>
<td>Sward, 2001</td>
<td>RCT</td>
<td>28 Combined (6M, 22F) 74</td>
<td>12 75 1 set 3 sessions</td>
</tr>
<tr>
<td>Tarnopolsky, 2007</td>
<td>non-RCT</td>
<td>8 Combined (10M, 3F) 75 M</td>
<td>24 75 3 sets 2 sessions</td>
</tr>
<tr>
<td>Tsutsuki, 1997</td>
<td>RCT</td>
<td>13 Combined (10M, 3F) 68</td>
<td>12 80 2 sets 3 sessions</td>
</tr>
<tr>
<td>Wells, 1995</td>
<td>non-RCT</td>
<td>5 Combined (11M, 3F) 67 M</td>
<td>12 70 3 sets 3 sessions</td>
</tr>
</tbody>
</table>