A meta-analysis of resistance training in female youth: its effect on muscular strength, and shortcomings in the literature

Running head: A meta-analysis of resistance training in female youth

Jason Moran[1]

Gavin Sandercock[2]

Rodrigo Ramirez-Campillo[3]

Cain C.T. Clark[1]

John F.T. Fernandes[1]

Benjamin Drury[1]

1. Department of Sport, University Centre Hartpury (University of the West of England), Gloucestershire, United Kingdom
2. Centre for Sports and Exercise Science, School of Biological Sciences, University of Essex, Colchester, United Kingdom,
3. Department of Physical Activity Sciences, Universidad de Los Lagos, Osorno, Chile

Corresponding author contact details: jason.moran@hartpury.ac.uk, +44 1452 702482

ABSTRACT

Background: Resistance training is an effective way to enhance strength in female youth but, to date, no researcher has meta-analysed its effect on muscular strength in that population.

Objectives: This meta-analysis characterised female youths' adaptability to resistance training (RT). A second objective was to highlight the limitations of the body of literature with a view to informing future research.

Data sources: Google Scholar, PubMed, Web of Science

Study eligibility criteria: Resistance training interventions in healthy females with a mean age between 8 and 18 years. Programmes of between 4 and 16 weeks duration that included a control group.

Study appraisal and synthesis methods: The inverse-variance random effects model for meta-analyses was used because it allocates a proportionate weight to trials based on the size of their individual standard errors and facilitates analysis whilst accounting for heterogeneity across studies. Effect sizes, calculated from a measure of muscular strength, are represented by the standardised mean difference and are presented alongside 95% confidence intervals.

Results: The magnitude of the main effect was 'small' (0.54, 95% confidence interval: 0.23, 0.85). Effect sizes were larger in older (> 15 yrs; ES = 0.72 [0.23, 1.21] vs. 0.38 [-0.02, 0.79]), taller (>163cm; ES = 0.67 [0.20, 1.13] vs. 0.55 [0.08, 1.02]) and heavier (<54kg; ES = 0.67 [0.30, 1.03] vs. 0.53 [-0.00, 1.06]) participants.

Conclusions and implications of key findings: Resistance training is effective in female youth. These findings can be used to inform the prescription of RT in female youth.

Key points:

- Resistance training is an effective way of increasing muscular strength in female youth.
- Older, taller or heavier female youths may be more responsive to training potentially owing to maturation-related increases in muscle mass.
- Programmes lasting 8 weeks, with 2 sessions per week and around 40 minutes per session seem most effective.

1. Introduction

Maximal strength is the maximum force or torque that can be exerted by skeletal muscles during movement [1]. The ability to exert high force is an important determinant of healthy function and athletic performance in youth [2] and resistance training (RT) is an effective way to enhance that quality [3] in this population. Indeed, previous research in youth has found that absolute strength is well correlated with certain measures of sprint and jump height (r = 0.596 to 0.762) [4,5].

Intervention studies of RT in female youth are lacking in both quantity and quality. Despite this, several meta-analytical reviews have investigated the efficacy of RT in youth populations [6–9]. However, to date, no researchers have conducted meta-analyses in female youth only, with existing data often conflated with that for young males [3,10] in subgroup analyses only, thus undermining the accuracy of inferences that can be made. This is an important matter in exercise science as there is variation in how male and female youth adapt to the demands of RT [11]. For example, adaptions in strength and body composition differ based on sex [11]. Primarily, this is because of differences in circulating anabolic hormones which are higher in males than in females from the age of puberty and which denote maturation-related changes such as increases in muscle mass [11,12]. This could result in an amplification of effects in male youth: males gain around 7.2kg of muscle mass annually during the growth spurt whilst females gain just 3.5kg per year [13]. On average, the female pubertal growth spurt takes place 2 years earlier than that for males but whilst gains in muscle mass slow down in females from the age of 15 years, males can continue to gain until the age of 20 years [13]. Accordingly, as maturation-status can regulate adaptive responses to RT [14], developmental factors must be considered in the prescription of this type of exercise.

The intertwining physiological processes that underpin increases in muscular strength have not been adequately detailed in line with the training methods that induce them in female youths. This primarily relates to factors such as sex and the specificity of training stimuli and though some authors have attempted to implicitly address such issues, reviews remain somewhat flawed. For example, Lesinski et al. [3] included studies in female youth athletes in their meta-analysis on RT but failed to separate them from male athletes for most effect estimates for chronological age, stage of maturation and training type. Indeed, dose responses were calculated independent of age and sex despite previous recommendations for age- and sex-specific approaches to RT prescription [15]. Behringer et al. [6] had previously reported that RT was effective for increasing strength in youth with an effect size of 1.1 (0.9-1.3). However, the main effect statistic was inclusive of both males and females. A later review by the same group [7] revealed beneficial effects of RT on the motor skills of running, jumping and throwing; and more recently, Harries, Lubans and Callister [9] meta-analysed the effects of RT on vertical jump in male and female youth athletes, reporting increases in vertical jump height in males during puberty, but not in females. Despite these encouraging results, the enhancement of motor skills may not necessarily be the primary effect associated with RT based on the principle of specificity [16] as such movements are heavily influenced by coordination [17] and body mass [18] whilst strength is defined by force production. In relation to this, little attention has been paid to the effects of RT on muscular strength only in female youth. Our aim was to address this gap in the literature by characterising the effects of RT on muscular strength in female youth, thus paying heed to the specificity of the response to training stimuli. A secondary aim was to inform future research and practice by describing shortcomings in the literature as it relates to RT in female youth.

2. Methods

This meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [19].

2.1 Literature search

With no date restrictions, a systematic search of the Google Scholar database was initially undertaken. Searches of the PubMed and Web of Science databases were also undertaken. Only articles published in the English language were considered. These searches were performed over three consecutive days (8^{th} to 10^{th}) in November, 2017. Using Boolean logic, various combinations of the following search terms were used: ‘youth’, ‘training’, ‘female’, ‘strength’, ‘resistance’, ‘weightlifting’, ‘volume’, ‘intensity’, ‘fitness’, ‘high’, ‘load’, ‘rest’, ‘sets’, ‘repetitions’, ‘plyometric’, ‘stretch-shortening cycle’, ‘jump’, ‘power’, ‘speed’, ‘velocity’, ‘agility’, ‘sprint’, ‘sprinting’, ‘alactic’, ‘acceleration’, ‘running’, ‘exercise’, ‘change of direction’, ‘paediatric’, ‘pediatric’, ‘young’, ‘children’, ‘adolescence’, ‘athletes’, ‘sport’. These combinations were searched using the following example format: ‘youth’ AND ‘training’ AND ‘female’ AND ‘strength’ OR ‘[additional search term 1]’ OR ‘[additional search term 2]’. In selecting studies for inclusion, a review of all relevant article titles was conducted before an examination of article abstracts and, then, full published articles. Only peer-reviewed articles were included in the meta-analysis. The reference lists of those studies that remained following the application of all inclusion criteria were hand-searched for further articles that could be relevant to the meta-analysis. Also, similar reviews were hand-searched for additional articles. The search process is outlined if Figure 1.

Figure 1 Flow chart for inclusion and exclusion of studies

2.2 Data extraction

The extraction of data from gathered articles was undertaken by three reviewers (JM, CC and JF) with a standardised form created in Microsoft Excel. The first reviewer collected the data before the second and third reviewers investigated its accuracy and the eligibility of studies for inclusion. Where required data were not clearly or completely reported, article authors were contacted for clarification.

The following criteria determined the eligibility of studies for inclusion in the review: cohorts of healthy females, with a mean age between 8 and 18 years. Interventions of between 4 and 16 weeks duration that included a control group. Based on a previous review [14], we defined RT as ““[requiring] the musculature to contract (sic.) against an opposing force generated by some type of resistance” [6]. Effect sizes were calculated by selecting the most relevant measure of muscular strength “based on theory or a logically defensible rationale” [20]. To account for the specificity of the training adaptation, we did not consider surrogate measures of performance such as a vertical jump [14] and tests of strength must have incorporated some form of external resistive load, and not bodyweight only. Means and standard deviations for a measure of post-intervention strength were used to calculate an effect size. The characteristics of the study participants are displayed in Table 1.

Table 1 Characteristics of study participants

2.3 Analysis and interpretation of results

Meta-analytical comparisons were carried out in RevMan version 5.3 [31]. The inverse-variance random effects model for meta-analyses was used because it allocates a proportionate weight to trials based on the size of their individual standard errors [32] and facilitates analysis whilst accounting for heterogeneity across studies [33]. Effect sizes are represented by the standardised mean difference and are presented alongside 95% confidence intervals. The calculated effect sizes were interpreted using the conventions outlined for standardised mean difference by Hopkins et al [34] (<0.2 = trivial; 0.2-0.6 = small, 0.6-1.2 = moderate, 1.2-2.0 = large, 2.0-4.0 = very large, >4.0 = extremely large). In cases in which there was more than one intervention group in a given study, the control group was proportionately divided to facilitate comparison across all participants [35].

To gauge the degree of heterogeneity amongst the included studies, the I^2 statistic was referred to. This represents the proportion of effects that are due to heterogeneity as opposed to chance [19]. Low, moderate and high levels of heterogeneity correspond to I^2 values of 25%, 50% and 75%, however, these thresholds are considered tentative [36]. The X^2 (chi square) statistic determines if the differences in the results of the analysis are due to chance and in such a case, a low P value, or high X^2 statistic, relative to degrees of freedom would be apparent [32].

A risk of bias quality scale was not utilised for a number of reasons: The Cochrane Collaboration has previously discouraged the use of these scales, stating that the practice is not underpinned by empirical evidence and assessment criteria may apply inaccurate study weights [37]. Also, the subjectivity of personal opinion undermines the accuracy of such scales [37]. Blinding of study participants and trainers is undermined owing to the constraints that make such a practice difficult to implement in training intervention studies [38]. In relation to this issue, previous systematic reviews [39,40] of training amongst children and adolescents suggest that studies tend to be of low to medium quality.

2.4 Analysis of moderator variables

To assess the potential effects of moderator variables, subgroup analyses were performed using moderating variables identified a priori. Using a random effects model we selected potential moderators likely to influence the effects of training on applied interventions. An age-based division was made between study groups whose participants were younger or older than 15 years. This division was made on the basis that this is the age at which females achieve adult height [41]. We also compared participants who were greater or less than 163 cm in stature. This represents an approximation of average female stature in the fully mature state [42]. Resistance training has previously been found to be less effective in pre-pubertal males in comparison to their pubertal and post-pubertal counterparts [14] and this was hypothesised to also be the case in females. With the high variability of body mass across various populations, body mass was divided into subgroups with a median split (greater or less than 56 kg) for the entire sample of participants. Age, height and body mass were all chosen as moderator variables because of the effect of age and biological maturation exert on performance in female youth [43]. The moderator variables of programme duration (weeks), training frequency (sessions per week), total number of training sessions and session duration (mins) were chosen based on the accepted influence of the FITT principle on adaptations to exercise [44]. These variables were divided using a median split.

3. Results

3.1 Main effect

Eleven studies were included in this meta-analysis and they comprised 15 individual experimental groups. Across all included studies, there was a small, significant improvement in strength (ES = 0.54 [0.23, 0.85], Z = 3.37 [p = 0.0008]). The overall estimate was of small magnitude and showed a significant level of between-study heterogeneity (I^2 = 42% [p = 0.04]). These results are displayed in Figure 2.

Figure 2 Forest plot of effect sizes with 95% confidence intervals

3.2 Effect of moderator variables

A summary of the effect of moderator variables can be viewed in Table 2. Subgroup analysis suggested highly variable levels of between-group heterogeneity. Older, taller and heavier study participants adapted to RT to a greater degree than their younger, shorter, lighter counterparts respectively with 'moderate' effects being achieved in the groups with larger adaptations. Interventions with a total amount of 16 sessions or fewer produced the largest effect (ES = 0.75 [0.33, 1.17], Z = 3.52 [P = 0.0004]). Those RT programmes that lasted 8 weeks or less demonstrated larger effects (ES = 0.62 [0.17, 1.07], Z = 2.71 [P = 0.007]) than those that lasted more than eight weeks (ES = 0.44 [-0.02, 0.90], Z = 1.86 [P = 0.06]). Similarly, interventions with two sessions per week were more effective (ES = 0.72 [0.34, 1.09], Z = 3.71 [P = 0.0002]) than those with more than 2 per week (ES = 0.18 [-0.26, 0.61], Z = 0.80 [P = 0.42]). Levels of heterogeneity were higher in subgroups with shorter programmes, lower training frequency, fewer training sessions and fewer minutes per session.

Table 2 Effect of moderator variables with 95% confidence intervals

4. Discussion

The purpose of this meta-analytical review was to quantify the effects of RT in female youth and to highlight limitations in the relevant body of literature with a view to improving approaches to research in the future. Based on extensive research findings in male youth

[14,45–49], the effects of RT in that population are well understood. The emergence of anabolic hormones in puberty [50] may enhance adaptations to RT due to the resultant increase in muscle mass [50]. However, because the magnitude of this adaptation is smaller in females [13], RT may be less effective in that population. This is evidenced by the results of this meta-analysis when compared to a methodologically similar meta-analysis of RT in male youth. Moran et al. [14] previously found a moderate effect (0.98 [0.70-1.27]) of RT on muscular strength in male youth aged 10 to 18 years. That contrasts with the current analysis which demonstrated effects of a far lower magnitude (0.54 [0.23, 0.85]). Given that it seems females adapt to RT to a different degree, this review is timely due to the dearth of empirical evidence in that population.

The body of research relating to RT in female youth is substantially smaller than that in male youth. A previous review of RT in male youth [14], which adopted similar study selection criteria, included 19 studies and excluded several more on the basis that it was a meta-analysis of within-group intervention effects and, thus, included no control trials or studies which included non-athletic youth. On the contrary, the current review found just 11 studies that met similar inclusion criteria, which were widened to include non-athletic youth. Given that RT can reduce injuries in females [15] and, also, that greater levels of strength are thought to prevent injury [51,52], it is curious that interventions which directly address the impact of RT on a measure of muscular strength are so scarce in female youth, a requirement for inclusion of interventions in this meta-analysis.

Based on the results of our analyses, the suggestion that RT can enhance strength in female youth is not in question but the next challenge from a research perspective is to characterise how maturation affects adaptations as the trainee develops physically. Previous meta-analyses [14,53,54] have reported variability in the degree of adaptation to various forms of training across the maturation continuum in male youth. However, the body of relevant literature is not currently large or informative enough to develop a stance on this issue in female youth. One of the primary causes of this is the relatively large amount of researchers

[55–57] who have incorporated RT into their studies but have not provided measures of absolute muscular strength with which to quantify the effects of their prescribed programmes (see Figure 1). This may, justifiably, be due to study design limitations but the approach could be argued to undermine the principle of specificity of adaptation to the imposed demands of training [58]. This is not necessarily a trivial issue: the making of inferences about effects on an arguably less-relevant outcome measure, such as vertical jump, due to a training method such as traditional RT, could be considered suboptimal due to the independent nature of training adaptations to different forms of training stimuli [16], and the principle of training specificity [58]. Our rationale for this stance is founded on the basis that, despite being well correlated with absolute strength in well-trained youth [4], measures of relative strength, such as vertical jump variants, may not necessarily capture maturation-related changes in strength [53]. This is because as youths grow, relative strength can decrease as bodyweight increases, potentially resulting in reduced relative strength, despite enhanced absolute strength [53]. This could be of particular concern in heavier youth. Though we appreciate that researchers' may have limited control over the type of test used, if possible, it is nonetheless important to consider this issue when formulating RT interventions and programmes so that the effects of training are measured with the most relevant performance tests. To measure performance following RT, researchers should consider utilising resistive apparatus which facilitates estimation of maximal strength to previously presented guidelines [59], in a variety of different exercises [60]. Such a protocol could be suitable for untrained youth as it involves the use of multiple submaximal repetitions to predict maximal strength [59]. The submaximal load used can be increased in manageable increments of between 2.5% and 5% until the participant exerts a maximal effort and whilst the study in question [59] used a submaximal resistance of 90%, lower loads could be justifiably prescribed.

In none of the studies included in this meta-analysis did authors report participants' maturation status with one of the most commonly used methods in youth sport [61]. This is a curious feature of the relevant body of literature given that altered motor coordination during puberty may influence females' susceptibility to damaging the anterior cruciate ligament, a risk that is deepened by deficits in strength which can predispose the knee to acute and chronic injury [62]. It is therefore logical to suggest that even an approximate estimate of the timing of peak height velocity could enable researchers to test training interventions that are specific to participants' level of biological maturation, thus targeting the aforementioned weaknesses. Given that in the majority of studies, researchers report participants' stature and body mass, the addition of seated stature to a typical testing battery is neither work-intensive or time-consuming. As recommended by the British Association of Sports and Exercise Sciences [63], researchers must report the biological maturity status in youth participants of both sexes.

Related to the above points, a further limitation of the current body of literature is the relatively high number of researchers who did not incorporate a control group into their study design. Several studies [64–67] were excluded on the basis that they did not provide any control group data, fulfilling other inclusion criteria. This study design feature takes on added significance in interventions in youth given that rapid changes in maturation status can result in both increases or decreases in physical capabilities [68,69]. The recruitment of individuals to studies can be difficult and the addition of a control group is not always possible. Nevertheless, with a view to progressing this area of research, study authors are encouraged to prioritise the inclusion of control groups during the conceptualisation stage of studies.

In reference to maturation-related factors, RT was more effective (moderate vs, small effects) in older, taller and heavier individuals, indicating that more biologically mature female youth may adapt to RT to a greater magnitude than their less mature counterparts. There also exists the possibility that potentially negative effects of puberty, such as impaired sensorimotor function [70], could have temporarily disrupted the progress of the less mature subgroups. This is in line with the suggestion that the existence of a maturational threshold could moderate responses to RT in youth [63]. Moreover, it is also reflective of a similar trend in male youth with the periods during and after peak height velocity seemingly an opportune time to expose well-conditioned individuals to more advanced and higher volume RT [14]. This may be partly due to maturation-related increases in muscle mass which, in turn, can underpin gains in strength [15]. Rising testosterone levels enhance the synthesis of muscle proteins which leads to the greater accumulation of muscle mass [71]. Though this process is less marked in female youth, who gain only half the amount of muscle mass that males do annually during the growth spurt [13], it still may exert a substantial effect. In line with the age threshold for subgroup analysis in this review, Poortmans et al. [72] found that 15 year old females possessed substantially more muscle mass than 10 year olds (18.4 ± 0.8 kg vs. 12.5 ± 0.4 kg) indicating a near-mature state for the older youths [73]. In females, the development of muscle mass during puberty seems to be related to a four-fold increase in testosterone which, though lower in magnitude than that experienced by males, still results in them achieving between 50% and 70% of the muscle mass of males [74]. Moreover, during the pubertal period, circulating oestrogen seems to promote fat storage, also resulting in lower muscle mass [71,75]. These factors together could go some way to explaining the size of the main effect in the current review which was almost half of that observed in a similar meta-analysis in male youth [14]. Indeed, whilst RT seems marginally more effective in older, taller and heavier individuals, unpublished data from our group seems to point to younger, shorter and lighter individuals being more responsive to plyometric training. This seems logical given that relative strength is a key factor in plyometric training whilst absolute strength is more important in RT. Taken together, this suggests that as in male youth, females may display variable responses to different types of training at different stages of maturation.

Contrary to what we had expected to observe from subgroup analyses, it seems that RT interventions were more effective in studies with fewer training sessions (16 or less), shorter study durations (8 weeks or less) and lower training frequencies (2 sessions or less per week). The reasons for this finding are not clear but could be due to the relatively low number of studies in this field, thus necessitating more research to clarify the time course of adaptation to RT in female youth. We surmise that the higher effects in shorter programmes could be indicative of a need to alter the training stimulus at the 8 week point in a training cycle: as evidenced by our results, adaptations to resistance training can take place in a relatively short period of time necessitating an eventual change in the demands placed on body in order to sustain progress [76]. This could also be indicative of females' lower sensitivity to the effects of RT when compared to males and could serve as a reason for coaches to plan more varied programmes of athletic development to continually drive adaptation in female youth. Despite this recommendation, the characteristics of the body of published literature could have been more influential in this finding: the lack of variation in programming characteristics such as session frequency made it somewhat difficult to divide studies with a median split for subgroup analysis. In relation to this, the dichotomisation of continuous data by median split can result residual confounding of analyses and reduced statistical power [77].

A widespread flaw in the literature relating to RT in female youth is the pooling of performance data of both females and males for analysis within the same studies [8,78–84]. In isolation, this is never an acceptable practice in research as it only determines whether a training method is effective independent of any population-specific effects. Such an approach does not consider the effects of sex and maturation level on training status given that boys and girls are biologically different and experience wholly different maturational changes at different times and tempos [85]. Given these differences, chronological age offers little basis for comparison between the sexes. Indeed, that researchers have often failed to provide separate within-study performance data for male and female youth renders many interventions' findings somewhat flawed and, arguably, reduces their usefulness to practitioners and sports scientists alike. To continue to drive progress in this area of research, study authors must report anthropometric and performance data of males and females separately, in addition to an analysis of an overall primary effect. For an exemplary demonstration of how to present quantitative and graphical data for boys and girls within a single intervention study, we refer the reader to the work Muehlbauer, Gollhoffer and Granacher [25].

5. Conclusion

Resistance training seems to be an effective way of increasing muscular strength in female youth. However, significant limitations in the current literature prevent assured RT prescription recommendations being made. Based on our results, it seems that older, taller or heavier individuals may be more responsive to training potentially owing to maturation-related increases in muscle mass which are indicative of a maturational threshold that could moderate responses to RT [63]. Following foundational training, females of any age can be exposed to RT though because adaptations may reach an upper limit relatively quicker than in boys, workloads should be sensibly balanced. This is particularly important during the interval of maximal growth when reduced motor control can result in injury. We recommend that two sessions per week is an adequate frequency of training in female youth who should be exposed to a varied training stimuli to prevent stagnation in the longer term. In relation to obtaining more relevant data for future analyses, researchers are urged to continue to investigate the effects of RT in female youth but to devote more effort into measuring maturation status and relevant measures of performance. If researchers include both males and females in the same study, the resultant data should be presented in a way that separates one group from the other so that sex-specific inferences can be made. Data should also be presented in its raw numerical form with graphical representations being used only to add context to the reported results.

Acknowledgements, conflicts of interest and funding statements

Jason Moran, Gavin Sandercock, Rodrigo Ramirez-Campillo, Cain C.T. Clark, John F.T. Fernandes and Benjamin Drury declare that they have no conflict of interest.

No financial support was received for the conduct of this study or preparation of this manuscript.

6. Reference list

1. Knuttgen HG, Komi P V. Considerations for Exercise. In: Komi P V, editor. Strength Power Sport. Oxford: Blackwell Science Ltd; 2003.

2. Faigenbaum AD, Kraemer WJ, Blimkie CJR, Jeffreys I, Micheli LJ, Nitka M, et al. Youth Resistance Training: Updated Position Statement Paper From the National Strength and Conditioning Association. J. Strength Cond. Res. 2009;23:S60–79.

3. Lesinski M, Prieske O, Granacher U. Effects and dose-response relationships of resistance training on physical performance in youth athletes: A systematic review and meta-analysis. Br. J. Sports Med. 2016. p. 781–95.

4. Comfort P, Stewart A, Bloom L, Clarkson B. Relationships between strength, sprint, and jump performance in well-trained youth soccer players. J. Strength Cond. Res. 2014;28:173–7.

5. Peñailillo L, Espíldora F, Jannas-Vela S, Mujika I, Zbinden-Foncea H. Muscle strength and speed performance in youth soccer players. J. Hum. Kinet. 2016;50:203–10.

6. Behringer M, Vom Heede A, Yue Z, Mester J. Effects of resistance training in children and adolescents: a meta-analysis. Pediatrics. 2010;126:e1199–210.

7. Behringer M, Vom Heede A, Matthews M, Mester J. Effects of strength training on motor performance skills in children and adolescents: a meta-analysis. Pediatr. Exerc. Sci. 2011;23:186–206.

8. Granacher U, Lesinski M, Büsch D, Muehlbauer T, Prieske O, Puta C, et al. Effects of resistance training in youth athletes on muscular fitness and athletic performance: A conceptual model for long-term athlete development. Front. Physiol. 2016.

9. Harries SK, Lubans DR, Callister R. Resistance training to improve power and sports performance in adolescent athletes: A systematic review and meta-analysis. J. Sci. Med. Sport. 2012. p. 532–40.

10. Payne VG, Morrow JR, Johnson L, Dalton SN. Resistance training in children and youth: a meta-analysis. Res. Q. Exerc. Sport [Internet]. 1997;68:80–8.

11. Yoshimoto T, Takai Y, Fukunaga Y, Fujita E, Yamamoto M, Kanehisa H. Effects of

school-based squat training in adolescent girls. J. Sports Med. Phys. Fitness. 2016;56:678–83.

12. Lloyd RS, Oliver JL. The Youth Physical Development Model: A new approach to Long-Term Athletic Development. Strength Cond. J. 2012;34:61–72.

13. Tonnessen E, Svendsen IS, Olsen IC, Guttormsen A, Haugen T. Performance Development in Adolescent Track and Field Athletes According to Age, Sex and Sport Discipline. PLoS One. 2015;10:e0129014.

14. Moran J, Sandercock GRH, Ramírez-Campillo R, Meylan C, Collison J, Parry DA. A meta-analysis of maturation-related variation in adolescent boy athletes' adaptations to short-term resistance training. J. Sports Sci. 2017;35:1041–51.

15. Lloyd RS, Faigenbaum AD, Stone MH, Oliver JL, Jeffreys I, Moody JA, et al. Position statement on youth resistance training: the 2014 International Consensus. Br. J. Sports Med. 2014;48:498–505.

16. Vissing K, Brink M, Lønbro S, Sørensen H, Overgaard K, Danborg K, et al. Muscle Adaptations to Plyometric vs. Resistance Training in Untrained Young Men. J. Strength Cond. Res. 2008;22:1799–810.

17. Hirtz P, Starosta W. Sensitive and Critical Periods of Motor Co-Ordination Development and Its Relation To Motor Learning. J. Hum. Kinet. 2002;7:19–28.

18. Kakebeeke TH, Lanzi S, Zysset AE, Arhab A, Messerli-Bürgy N, Stuelb K, et al. Association between body composition and motor performance in preschool children. Obes. Facts. 2017;10:420–31.

19. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J. Clin. Epidemiol. 2009. p. e1-34.

20. Turner HM, Bernard RM. Calculating and synthesizing effect sizes. Contemp. Issues Commun. Sci. Disord. 2006;33:42–55.

21. Eather N, Morgan PJ ame., Lubans DR evald. Improving health-related fitness in adolescents: the CrossFit Teens™ randomised controlled trial. J. Sports Sci. 2016;34:209–23.

22. Ignjatovic AM, Markovic ZM, Radovanovic DS. Effects of 12-week medicine ball training on muscle strength and power in young female handball players. J. Strength Cond. Res. 2012;26:2166–73.

23. Johnson AW, Eastman CS, Feland JB, Mitchell UH, Mortensen BB, Eggett D. Effect of high-speed treadmill training with a body weight support system in a sport acceleration program with female soccer players. J. Strength Cond. Res. 2013;27:1496–502.

24. Lillegard WA, Brown EW, Wilson DJ, Henderson R, Lewis E. Efficacy of strength training in prepubescent to early postpubescent males and females: Effects of gender and maturity. Dev. Neurorehabil. 1997;1:147–57.

25. Muehlbauer T, Gollhofer A, Granacher U. Sex-Related Effects in Strength Training during Adolescence: A Pilot Study. Percept. Mot. Skills. 2012;115:953–68.

26. Pereira A, Costa AM, Santos P, Figueiredo T, João PV. Training strategy of explosive strength in young female volleyball players. Medicina (B. Aires). 2015;51:126–31.

27. Santos A, Marinho DA, Costa AM, Marques MC, Santos A, Marinho DA, et al. The Effects of Concurrent Resistance and Endurance Training Follow a Specific Detraining Cycle in Young School Girls. J. Hum. Kinet. 2011;93–103.

28. Siegel JA, Camaione DN, Manfredi TG. The Effects of Upper Body Resistance Training on Prepubescent Children. Pediatr. Exerc. Sci. 1989;1:1451–4.

29. Skattebo Ø, Hallén J, Rønnestad BR, Losnegard T. Upper body heavy strength training does not affect performance in junior female cross-country skiers. Scand. J. Med. Sci.

Sports. 2015;n/a-n/a.

30. Sugimoto D, Myer GD, Bush HM, Hewett TE. Effects of compliance on trunk and hip integrative neuromuscular training on hip abductor strength in female athletes. J. Strength Cond. Res. 2014;28:1187–94.

31. The Nordic Cochrane Centre. Review Manager. Cochrane Collab. 2014. p. 1–43.

32. Deeks JJ, Higgins JP, Altman DG. Analysing Data and Undertaking Meta-Analyses. Cochrane Handb. Syst. Rev. Interv. Cochrane B. Ser. 2008. p. 243–96.

33. Kontopantelis E, Springate DA, Reeves D. A Re-Analysis of the Cochrane Library Data: The Dangers of Unobserved Heterogeneity in Meta-Analyses. PLoS One. 2013;8.

34. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. Med. Sci. Sports Exerc. 2009. p. 3–12.

35. Higgins JP, Deeks JJ, Altman DG. Special Topics in Statistics. Cochrane Handb. Syst. Rev. Interv. Cochrane B. Ser. 2008. p. 481–529.

36. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ Br. Med. J. 2003;327:557–60.

37. Higgins JP, Altman DG. Assessing Risk of Bias in Included Studies. Cochrane Handb. Syst. Rev. Interv. Cochrane B. Ser. 2008. p. 187–241.

38. Bolger R, Lyons M, Harrison AJ, Kenny IC. Sprinting Performance and Resistance-Based Training Interventions. J. Strength Cond. Res. 2015;29:1146–56.

39. Johnson BA, Salzberg CL, Stevenson DA. A Systematic Review: Plyometric Training Programs for Young Children. J. Strength Cond. Res. 2011;25:2623–33.

40. Bedoya A, Miltenberger MR, Lopez RM. Plyometric Training Effects on Athletic Performance in Youth Soccer Athletes. J. Strength Cond. Res. 2015.

41. Georgopoulos NA, Markou KB, Theodoropoulou A, Vagenakis GA, Benardot D, Leglise

M, et al. Height velocity and skeletal maturation in elite female rhythmic gymnasts. J. Clin. Endocrinol. Metab. 2001;86:5159–64.

42. Pellett PL. Food energy requirements in humans. Am. J. Clin. Nutr. 1990. p. 711–22.

43. Emmonds S, Morris R, Murray E, Robinson C, Turner L, Jones B. The influence of age and maturity status on the maximum and explosive strength characteristics of elite youth female soccer players. Sci. Med. Footb. 2017;1:209–15.

44. Pescatello LS, MacDonald H V., Lamberti L, Johnson BT. Exercise for Hypertension: A Prescription Update Integrating Existing Recommendations with Emerging Research. Curr. Hypertens. Rep. 2015.

45. Radnor JM, Lloyd RS, Oliver JL. Individual Response to Different Forms of Resistance Training in School-Aged Boys. J. Strength Cond. Res. 2017;31:787–97.

46. Lloyd R, Radnor J, De Ste Croix M, Cronin J, Oliver J. Changes in sprint and jump performances after traditional, plyometric and combined resistance training in male youth pre and post peak height velocity. J. Strength Cond. Res. 2016;30:1239–1247.

47. Moran J, Sandercock GRH, Ramírez-Campillo R, Wooller J-J, Logothetis S, Schoenmakers PPJM, et al. Maturation-related differences in adaptations to resistance training in young male swimmers. J. strength Cond. Res. 2018; 32:139-149.

48. Rumpf MC, Cronin JB, Mohamad IN, Mohamad S, Oliver JJL, Hughes MG. The effect of resisted sprint training on maximum sprint kinetics and kinematics in youth. Eur. J. Sport Sci. 2015;15:374–81.

49. Meylan CMP, Cronin JB, Oliver JL, Hopkins WG, Contreras B. The effect of maturation on adaptations to strength training and detraining in 11-15-year-olds. Scand. J. Med. Sci. Sport. 2014;24:156–64.

50. Viru A, Loko J, Harro M, Volver A, Laaneots L, Viru M. Critical Periods in the Development of Performance Capacity During Childhood and Adolescence. Eur. J. Phys.

Educ. 1999;4:75–119.

51. Faigenbaum AD, Myer GD. Resistance training among young athletes: safety, efficacy and injury prevention effects. Br. J. Sports Med. 2010;44:56–63.

52. Askling C, Karlsson J, Thorstensson A. Hamstring injury occurrence in elite soccer players after preseason strength training with eccentric overload. Scand. J. Med. Sci. Sport. 2003;13:244–50.

53. Moran J, Sandercock G, Rumpf MC, Parry DA. Variation in Responses to Sprint Training in Male Youth Athletes: A Meta-analysis. Int. J. Sports Med. 2017;38:1–11.

54. Moran J, Sandercock GRH, Ramírez-Campillo R, Meylan C, Collison J, Parry DA. Age-related variation in male youth athletes' countermovement jump following plyometric training. J. Strength Cond. Res. 2017;31:552–65.

55. Siegler J, Gaskill S, Ruby B. Changes evaluated in soccer-specific power endurance either with or without a 10-week, in-season, intermittent, high-intensity training protocol. J. Strength Cond. Res. 2003;17:379–87.

56. Sannicandro I, Cofano G, Rosa RA, Piccinno A. Balance training exercises decrease lower-limb strength asymmetry in young tennis players. J. Sport. Sci. Med. 2014;13:397–402.

57. Pairot De Fontenay B, Lebon F, Champely S, Argaud S, Monteil1 K. ACL Injury Risk Factors Decrease & Jumping Performance Improvement in Female Basketball Players: A Prospective Study. Int. J. Kinesiol. Sport. Sci. 2013;1:2201–6015.

58. Silva JM. An analysis of the training stress syndrome in competitive athletics. J. Appl. Sport Psychol. 1990;2:5–20.

59. Verdijk LB, Van Loon L, Meijer K, Savelberg HHCM. One-repetition maximum strength test represents a valid means to assess leg strength in vivo in humans. J. Sports Sci. 2009;27:59–68.

60. Faigenbaum AD, Milliken L a, Westcott WL. Maximal strength testing in healthy children. J. Strength Cond. Res. 2003;17:162–6.

61. Mirwald RL, Baxter-Jones ADG, Bailey DA, Beunen GP. An assessment of maturity from anthropometric measurements. Med. Sci. Sports Exerc. 2002;34:689–94.

62. Quatman-Yates CC, Myer GD, Ford KR, Hewett TE. A longitudinal evaluation of maturational effects on lower extremity strength in female adolescent athletes. Pediatr. Phys. Ther. 2013;25:271–6.

63. McNarry M, Lloyd RS, Buchheit M, Craig P, Fbases W, Oliver J. The BASES Expert Statement on Trainability during Childhood and Adolescence. Sport Exerc. Sci. 2014;22–3.

64. Naczk M, Lopacinski A, Brzenczek-Owczarzak W, Arlet J, Naczk A, Adach Z. Influence of short-term inertial training on swimming performance in young swimmers. Eur. J. Sport Sci. 2017;17:369–77.

65. Dowse RA, McGuigan MR, Harrison C. Effects of a Resistance Training Intervention on Strength, Power, and Performance in Adolescent Dancers. J. Strength Cond. Res. 2017;In press.

66. Ullrich B, Pelzer T, Oliveira S, Pfeiffer M. Neuromuscular Responses to Short-Term Resistance Training with Traditional and Daily Undulating Periodization in Adolescent Elite Judoka. J. Strength Cond. Res. 2016;30:2083–99.

67. Steele J, Fisher JP, Assunção AR, Bottaro M, Gentil P. The role of volume-load in strength and absolute endurance adaptations in adolescent’s performing high- or low-load resistance training. Appl. Physiol. Nutr. Metab. 2017;42:193–201.

68. Hewett TE, Myer GD, Ford KR. Decrease in neuromuscular control about the knee with maturation in female athletes. J. Bone Joint Surg. Am. 2004;86–A:1601–8.

69. Yagüe PH, La Fuente D, Manuel J. Changes in height and motor performance relative to peak height velocity: A mixed-longitudinal study of Spanish boys and girls. Am. J. Hum. Biol.

1998;10:647–60.

70. Quatman-Yates CC, Quatman CE, Meszaros AJ, Paterno M V., Hewett TE. A systematic review of sensorimotor function during adolescence: A developmental stage of increased motor awkwardness? Br. J. Sports Med. 2012. p. 649–55.

71. Tipton KD. Gender differences in protein metabolism. Curr. Opin. Clin. Nutr. Metab. Care. 2001;4:493–8.

72. Poortmans JR, Boisseau N, Moraine JJ, Moreno-Reyes R, Goldman S. Estimation of total-body skeletal muscle mass in children and adolescents. Med. Sci. Sports Exerc. 2005;37:316–22.

73. Siervogel RM, Demerath EW, Schubert C, Remsberg KE, Chumlea WC, Sun S, et al. Puberty and body composition. Horm. Res. 2003. p. 36–45.

74. Welsman J, Armstrong N, Kirby B. Children and Exercise XiX. Washington: Singer; 1997.

75. Garnett SP, Högler W, Blades B, Baur LA, Peat J, Lee J, et al. Relation between hormones and body composition, including bone, in prepubertal children. Am. J. Clin. Nutr. 2004;80:966–72.

76. Kraemer WJ, Ratamess NA. Fundamentals of Resistance Training: Progression and Exercise Prescription. Med. Sci. Sports Exerc. 2004. p. 674–88.

77. Altman D, Royston P. The cost of dichotomising continuous variables. BMJ. 2006;332:1080.

78. Assunção AR, Bottaro M, Ferreira-Junior JB, Izquierdo M, Cadore EL, Gentil P. The chronic effects of low- and high-intensity resistance training on muscular fitness in adolescents. PLoS One. 2016;11.

79. Benson AC, Torode ME, Fiatarone Singh MA. The effect of high-intensity progressive

resistance training on adiposity in children: a randomized controlled trial. Int J Obes. 2008;32:1016–27.

80. Faigenbaum AD, Westcott2 WL, Micheli LJ, Outerbridge a. R, Long CJ, LaRosa-Loud R, et al. The Effects of Strength Training and Detraining on Children. J. Strength Cond. Res. 1996;10:109.

81. Velez A, Golem DL, Arent SM. The impact of a 12-week resistance training program on strength, body composition, and self-concept of hispanic adolescents. J. Strength Cond. Res. 2010;24:1065–73.

82. Faigenbaum AD, Zaichkowsky LD, Westcott WL, Micheii LJ, Fehlandt AF. The Effects of a Twice-a-Week Strength Training Program on Children. Pediatr. Exerc. Sci. 1993;5:339–46.

83. Faigenbaum AD, Milliken LA, Loud RL, Burak BT, Doherty CL, Westcott WL. Comparison of 1 and 2 Days per Week of Strength Training in Children. Res. Q. Exerc. Sport. 2002;73:416–24.

84. Faigenbaum AD, Milliken L, Moulton L, Westcott W. Early Muscular Fitness Adaptations In Children In Response To Two Different Resistance Training Regimens. Med. Sci. Sport. Exerc. 2005;37:S185.

85. Marceau K, Ram N, Houts RM, Grimm KJ, Susman EJ. Individual differences in boys' and girls' timing and tempo of puberty: Modeling development with nonlinear growth models. Dev. Psychol. 2011;47:1389–409.

Table

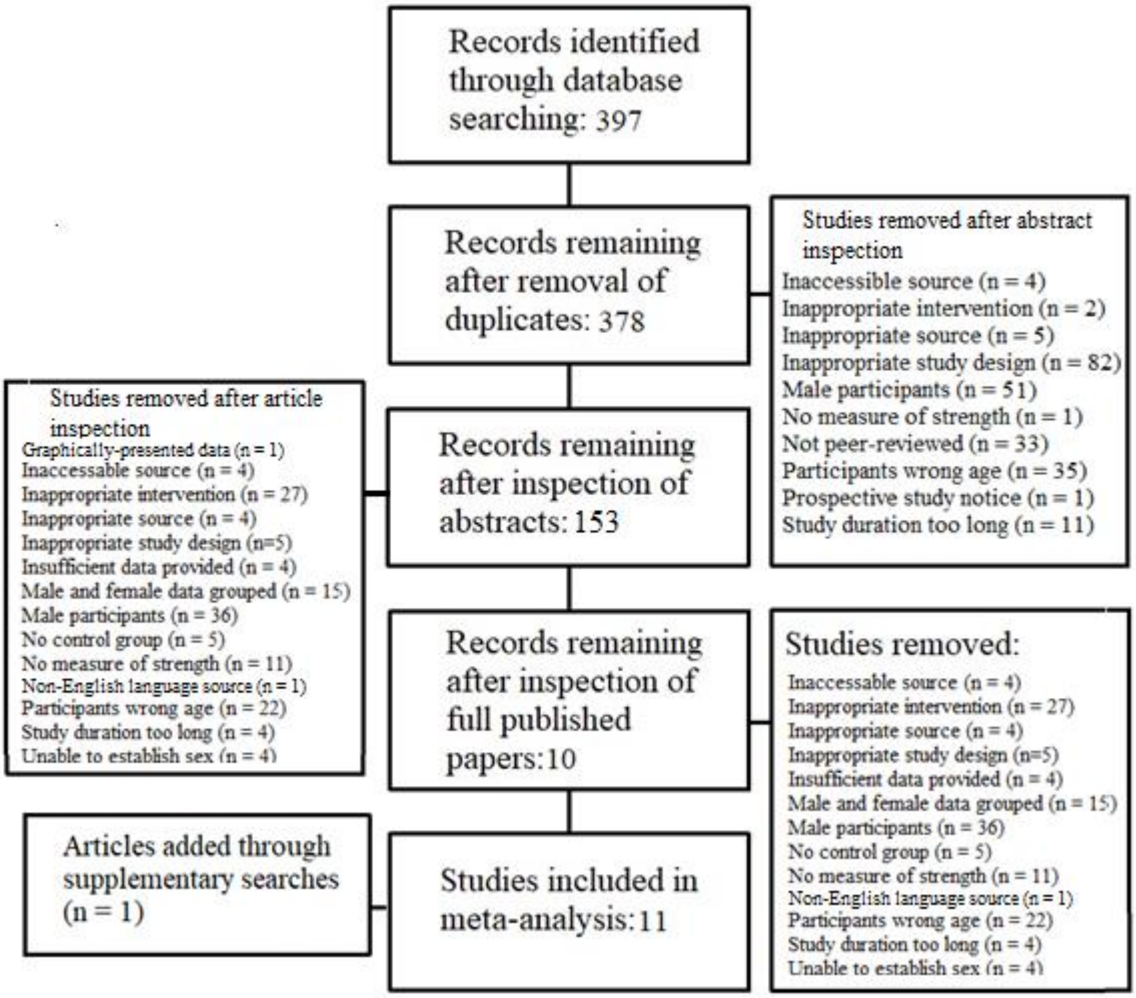
Click here to download Table Table 1.docx

Study	Study group	Sport	Age	Height (cm)	Weight	Participants	Weeks	Frequency (per week)	Total sessions	Session duration (mins)	Test
Eather et al. [21]	Crossfit Teens	Physical education	15.5	165	58.5	31	8	2	16	60	Grip-strength (kg)
	Control group	Physical education	15.5	165	58.5	19					Grip-strength (kg)
Ignjatovic et al. [22]	Experimental group	Handball	16.9			11	12	2	24	15	1 repetition maximum bench press (kg)
	Control group	Handball	16.9			10					1 repetition maximum bench press (kg)
Johnson et al. [23]	High-speed treadmill (HST)	Soccer	16	164	58.1	12	6	2	12		Isometric knee flexor strength (kg) (leg extension)
	Standard treadmill (ST)	Soccer	16	164	58.1	12	6	2	12		Isometric knee flexor strength (kg) (leg extension)
	Control group	Soccer	16	164	58.1	8					Isometric knee flexor strength (kg) (leg extension)
Lillegard et al. [24]	Tanner stage 1-2 (Experimental)	General population	9.5	136.3	36.7	7	12	3	36	40	10RM leg extension (kg)
	Tanner stage 1-2 (Control)	General population	9.6	137.3	35.3	5					10RM leg extension (kg)
	Tanner stage 3-5 (Experimental)	General population	13.2	158.1	50	6	12	3	36	40	10RM leg extension (kg)
	Tanner stage 3-5 (Control)	General population	12.5	153.3	50.3	4					10RM leg extension (kg)
Muehlbauer et al. [25]	Strength training group	Various	16.6	161.5	55.1	8	8	2	16		Maximum isometric force (N) (leg press)
	Control group	Various	16.7	166	54.4	7					Maximum isometric force (N) (leg press)
Pereira et al. [26]	Experimental group	Volleyball	14	160	52	10	8	2	16		1.5Kg Medicine ball throwing (m)
	Control group	Volleyball	13.8	160	53.5	10					1.5Kg Medicine ball throwing (m)
Santos et al. [27]	Combined strength and endurance training (GCOM)	Various	13.5	157.9	54.8	25	8	2	16		3Kg Medicine ball throwing (m)
	Strength training (GR)	Various	13.5	159.4	58.9	21	8	2	16		3Kg Medicine ball throwing (m)
	Control	Various	13.5	156.8	51.5	21					3Kg Medicine ball throwing (m)
Siegel et al. [28]	Experimental group	General population	8.5	128.5	27.2	24	12	3	36	30	Cable flexion (kg)
	Control group	General population	8.4	128.2	26.4	16					Cable flexion (kg)
Skattebo et al. [29]	Intervention group	Cross-country skiing	18	171	61	9	10	2	20	33	1RM seated pull-down (kg)
	Control Group	Cross-country skiing	17	166	60	7					1RM seated pull-down (kg)

Sugimoto et al. [30]	Moderate-compliance group	Volleyball	15.5	170.3	59.4	7	10	3	30	Hip abductor peak torque (ft-lbs) (isokinetic dynamometer)
	High-compliance group	Volleyball	15.5	171.6	68.8	7	10	3	30	Hip abductor peak torque (ft-lbs) (isokinetic dynamometer)
	Noncompliance group	Volleyball	16	173.4	63.9	7				Hip abductor peak torque (ft-lbs) (isokinetic dynamometer)
Yoshimoto et al. [11]	Training group	Various	13.8	154.9	48.2	27	8	5.5	44	Maximal knee extension (kg)
	Control group	Various	13.8	154.4	46.4	20				Maximal knee extension (kg)

Subgroup	Effect size and confidence interval	Effect descriptor	*P*	Groups	N	Between group I^2	Between group *P*	Within group I^2	Within group *P*
< 15yrs	0.38 [-0.02, 0.79]	Small	0.06	7	196	7.6%	0.30	41%	0.12
> 15yrs	0.72 [0.23, 1.21]	Moderate	0.004	8	155			42%	0.10
< 163cm	0.55 [0.08, 1.02]	Small	0.02	8	211	0%	0.73	57%	0.02
> 163cm	0.67 [0.20, 1.13]	Moderate	0.005	6	119			18%	0.30
< 56kg	0.53 [-0.00, 1.06]	Small	0.05	7	180	0%	0.68	60%	0.02
> 56kg	0.67 [0.30, 1.03]	Moderate	0.0004	7	150			3%	0.40
≤ 8wks	0.62 [0.17, 1.07]	Moderate	0.007	8	231	0%	0.58	56%	0.03
> 8wks	0.44 [-0.02, 0.90]	Small	0.06	7	120			25%	0.23
≤ 2 sess p/week	0.72 [0.34, 1.09]	Moderate	0.0002	9	221	70.1%	0.07	36%	0.13
> than 2 sess p/week	0.18 [-0.26, 0.61]	Trivial	0.42	6	130			21%	0.27
≤16 sessions	0.75 [0.33, 1.17]	Moderate	0.0004	7	184	54.6%	0.14	35%	0.16
> 16 sessions	0.30 [-0.11, 0.72]	Small	0.16	8	167			33%	0.16
< 40 mins per session	0.34 [-0.38, 1.06]	Small	0.03	3	77	0%	0.52	53%	0.12
≥ 40 mins per session	0.64 [0.11, 1.16]	Moderate	0.02	4	72			4%	0.35

Table 2 Effect of moderator variables with 95% confidence intervals


Records identified through database searching: 397
Records remaining after removal of duplicates: 378
Studies removed after abstract inspection
Inaccessible source (n = 4)
Inappropriate intervention (n = 2)
Inappropriate source (n = 5)
Inappropriate study design (n = 82)
Male participants (n = 51)
No measure of strength (n = 1)
Not peer-reviewed (n = 33)
Participants wrong age (n = 35)
Prospective study notice (n = 1)
Study duration too long (n = 11)
Studies removed after article inspection
Graphically-presented data (n = 1)
Inaccessable source (n = 4)
Inappropriate intervention (n = 27)
Inappropriate source (n = 4)
Inappropriate study design (n=5)
Insufficient data provided (n = 4)
Male and female data grouped (n = 15)
Male participants (n = 36)
No control group (n = 5)
No measure of strength (n = 11)
Non-English language source (n = 1)
Participants wrong age (n = 22)
Study duration too long (n = 4)
Unable to establish sex (n = 4)
Records remaining after inspection of abstracts: 153
Records remaining after inspection of full published papers: 10
Studies removed:
Inaccessable source (n = 4)
Inappropriate intervention (n = 27)
Inappropriate source (n = 4)
Inappropriate study design (n=5)
Insufficient data provided (n = 4)
Male and female data grouped (n = 15)
Male participants (n = 36)
No control group (n = 5)
No measure of strength (n = 11)
Non-English language source (n = 1)
Participants wrong age (n = 22)
Study duration too long (n = 4)
Unable to establish sex (n = 4)
Articles added through supplementary searches (n = 1)
Studies included in meta-analysis: 11

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
Eather et al. [21]	30.5	8.9	31	26.8	8.9	19	10.9%	0.41 [-0.17, 0.99]
Ignjatovic et al. [22]	40.1	6.8	11	39.8	4.7	10	7.6%	0.05 [-0.81, 0.91]
Johnson et al. [23] (HST)	31.55	8.9	12	24.76	7.4	4	5.1%	0.75 [-0.42, 1.92]
Johnson et al. [23] (ST)	38.01	7.1	12	24.76	7.4	4	4.2%	1.75 [0.42, 3.08]
Lillegard et al. [24] (Tanner 1-2)	19.3	9.6	7	10.4	4.7	5	4.6%	1.03 [-0.23, 2.28]
Lillegard et al. [24] (Tanner 3-5)	37.4	7.3	6	23.8	10.1	4	3.5%	1.45 [-0.05, 2.96]
Muehlbauer et al. [25]	1,639	324.5	8	1,061.9	198.3	7	4.3%	1.99 [0.68, 3.29]
Pereira et al. [26]	7.9	14.3	10	6	43.8	10	7.4%	0.06 [-0.82, 0.93]
Santos et al. [27] (GCOM)	4.25	0.73	25	3.76	0.43	11	8.9%	0.73 [-0.00, 1.46]
Santos et al. [27] (GR)	4.29	0.74	21	3.76	0.43	10	8.3%	0.78 [0.00, 1.56]
Siegel et al. [28]	11.8	1.9	24	11.8	2	16	10.1%	0.00 [-0.63, 0.63]
Skattebo et al. [29]	34.67	3.12	9	29.29	4.68	7	5.4%	1.32 [0.20, 2.43]
Sugimoto et al. [30] (High compliance)	55.4	13.1	7	51.6	6.1	3	4.1%	0.29 [-1.07, 1.65]
Sugimoto et al. [30] (Moderate compliance)	51.4	7.8	7	51.6	6.1	4	4.7%	-0.03 [-1.25, 1.20]
Yoshimoto et al. [11]	39.8	9.8	27	41.5	7.4	20	10.9%	-0.19 [-0.77, 0.39]
Total (95% CI)			**217**			**134**	**100.0%**	**0.54 [0.23, 0.85]**

Heterogeneity: Tau² = 0.15; Chi² = 24.13, df = 14 (P = 0.04); I² = 42%

Test for overall effect: Z = 3.37 (P = 0.0008)