Effects of resistance training interventions on muscular strength in adults with intellectual disability: a systematic review and meta-analysis

Iva Obrusnikova[a], Cora J. Firkin[a], Albert R. Cavalier[b] and Richard R. Suminski[a]

[a]Department of Behavioral Health and Nutrition, University of Delaware, Newark, USA;

[b]School of Education, University of Delaware, Newark, USA

Abstract

Purpose: Muscular strength is critical for adults with an intellectual disability (ID) to promote their mobil-ity, cardiovascular capacity, and performance of daily living/recreational/vocational activities. This article reports the results of the first systematic review and meta-analysis of peer-reviewed clinical trials that evaluated the effects of resistance training (RT) interventions on muscular strength in adults with ID. Methods: The protocol was registered with PROSPERO (CRD42020184905). The review focuses on clinical trials that recorded quantitative measures of maximum muscular strength. Eleven electronic databases were searched from their earliest available record up to May 2020. After screening 1996 search records, 11 clinical trials were reviewed.
Results: The RT interventions, while heterogeneous, had an overall significant ($p \leq 0.05$) effect on muscu-lar strength in adults with ID, ages 25–58 years. The findings were more significant and less heteroge-neous for non-combined RT interventions than for interventions that combined RT exercises with aerobic or balance exercises. The TESTEX overall score was 8.3 ± 3.6.
Conclusions: RT interventions (particularly when not combined with other exercises) are effective in pro-moting muscular strength in adults with ID. The limited number of studies and the low study quality scores indicate a potential risk of bias, which limits the interpretation of the findings and warrants further investigation

Implications for Rehabilitation

- Muscular strength is critical for adults with an intellectual disability (ID) to promote their mobility, cardiovascular capacity, and performance of daily living/recreational/vocational activities.

- RT interventions are an effective means of improving muscular strength in adults with ID, especially when not combined with other forms of exercise.

- Testing and assessment protocols used in RT programs should be individualized for adults with ID to accommodate their characteristics and should be implemented under conditions similar to those experienced during the training regimen.

- It is important to implement familiarization sessions before carrying out muscular strength testing or initiating an RT program to ensure safety, accuracy, and effectiveness of the program for adults with ID.

Introduction

Adequate physical activity (PA) levels play a significant role in overall health status and chronic disease prevention and manage-ment for adults with an intellectual disability (ID) [1,2]. A recent review [3] revealed that 91% of adults with ID do *not* meet the *Physical Activity Guidelines for Americans* (PAG) [4]. PAG recom-mends that adults participate in at least 150–300 min of moder-ate-intensity aerobic PA per week, or 75–150 min of vigorous-intensity PA per week. Adults also should participate in muscle-strengthening activities of moderate or greater intensity that involve all major muscle groups on at least two days per week. Though much remains to be learned about the benefits of PA for specific types of disabilities, sufficient evidence exists to

recommend that adults with disabilities should avoid inactivity and participate in regular PA according to their abilities [4].

One approach to meeting PAG is incorporating regular exercise in a person's life. Both PAG and the World Health Organization [5] identify exercise as a subcategory of PA that is planned, struc-tured, repetitive, and designed to improve or maintain one or more components of physical fitness. PAG [4] defines physical fit-ness as "the ability to carry out daily tasks with vigor and alert-ness, without undue fatigue, and with ample energy to enjoy leisure-time pursuits and respond to emergencies" (p. 33). Physical fitness components are classified into two distinct cate-gories: (1) health-related fitness (HRF) components (cardiorespira-tory fitness, muscular strength, muscular endurance, body composition, and flexibility) and (2) skill- or performance-related

fitness components (balance, reaction time, coordination, agility, speed, and power) [6,7]. Promoting HRF components is especially important in adults with ID. Many systematic reviews have revealed that, compared to the general population, adults with ID exhibit lower levels of cardiovascular fitness [8,9], muscular strength and endurance [8,10,11], and a higher prevalence of obesity [12,13]. Studies also have revealed that adults with ID experience earlier aging [14], loss of independence [15], and a higher prevalence of chronic diseases [16]. Increased levels of physical fitness and PA are critical to the prevention of the onset of these negative health consequences and risk factors [17,18].

Muscular strength is a component of HRF. The American College of Sports Medicine (ACSM) [6] defines muscular strength as "the ability of a muscle group to develop maximal contractile force against a resistance in a single contraction, and is related to the ability to perform activities that require high levels of muscular force" (p. 3). ACSM [6] divides muscle actions into two basic types: isometric/static and dynamic. During an *isometric action*, the muscle generates force without movement taking place (e.g., pushing or pulling against an immovable object or holding an object in place). Isometric actions traditionally are assessed with devices specific to a muscle group and joint angle of testing, such as a handgrip dynamometer and a cable tensiometer [6,19]. During a *dynamic action*, the muscle generates force to move an object during which the muscle changes in length. The changes in muscle length can be either eccentric (i.e., the muscle lengthens) or concentric (i.e., the muscle shortens). Additionally, dynamic actions can be either isotonic (i.e., involving a fixed amount of resistance) or isokinetic (i.e., involving a fixed speed). Dynamic contractions traditionally are assessed by a one-repetition maximum (1-RM), in which a person exerts the greatest force through a full range of motion or by an isokinetic muscular performance test, which uses an isokinetic dynamometer, such as the Biodex or Cybex. These computerized dynamometers measure the force or moment exerted by the muscle or muscle groups contracting against a controlled accommodating resistance [6,19]. Static and dynamic actions are physiologically different indications of muscular strength [20]. Even though they correlate significantly, muscles adapt differently to static and dynamic training regimens. Therefore, testing should involve conditions similar to those experienced by the person during the training regimen, in terms of the structure of the test, the mode and velocity of contraction, and the load or resistance [21].

The health benefits of enhanced muscular strength in the general population are well established [22]. In adults with ID, increased levels of muscular strength are positively associated with an improved cardiovascular capacity [23] and performance of recreational or vocational activities and activities of daily living [23–25]. Low levels of muscular strength, however, are correlated with body sway and functional decline [26,27]. Resistance training (RT) interventions aimed at increasing muscular strength in adults with ID have increasingly attracted researchers and clinicians over the past two decades [28,29]. RT is defined by ACSM [30] as a form of strength training that is designed to improve muscular strength, power, and endurance. It involves the activation of motor units against external resistance that can be applied to whole-body movements or isolated muscle groups. A range of equipment can be used to apply external resistance (e.g., body-weight, free weights, machines with additional weights, elastic bands, or water pressure). An RT program is designed for a specific person by adjusting acute training variables, such as the choice and order of exercises, frequency of exercise sessions, a number of sets and repetitions, intensity levels, and duration of

rest periods [30]. Studies with adults with ID found that some RT interventions have improved muscular strength and endurance [28], functional performance [31], vocational and athletic performance [32,33], and levels of PA [34].

While a number of reviews have broadly synthesized and evaluated the effects of various exercise interventions on HRF components in adults with ID [18,28,29,35], none of them specifically evaluated the effects of RT interventions on muscular strength in adults with ID. Further, the reviews identified methodological problems in the HRF research studies that include lack of randomization and allocation concealment, use of instruments and procedures with inadequate measurement properties, and lack of participants' familiarization with the test procedure or exercises. As a result, the mechanisms by which persons with ID improve their muscular strength and derive their associated benefits remain poorly understood, thereby reducing the likelihood of establishing a coherent and integrated body of knowledge upon which future investigations and interventions can be based.

Therefore, the primary aim of this article is to perform a systematic review and meta-analysis of published, peer-reviewed clinical trials of RT interventions to evaluate their effects on muscular strength in adults with ID. We hypothesized that RT interventions (whether alone or combined with other exercises) have significant effects ($p \leq 0.05$) on measures of muscular strength in adults with ID. Secondary aims are to (1) summarize the study designs and participant characteristics; (2) summarize the FITT characteristics of the RT interventions; (3) assess the validity and reliability of the main measures of maximum muscular strength; (4) evaluate the effects of different RT interventions on maximum muscular strength in adults with ID; (5) evaluate the risk of bias; and (6) provide recommendations on how to improve the quality of future research. The findings of this review and meta-analysis will inform researchers in the development of evidence-based RT programs that will effectively promote muscular strength in adults with ID. Furthermore, the findings will help establish evidence-based guidelines for health professionals in prescribing RT to adults with ID.

Method

Protocol registration and search strategy

The protocol for this systematic review was published online at the International Prospective Register of Systematic Review (PROSPERO) on May 7th, 2020 (CRD42020184905). The review was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [36]. A search of English-language, peer-reviewed articles was performed using the following major databases: ERIC, MEDLINE, PsycINFO, and Sports Medicine and Education Index *via* Proquest, ERIC, CINAHL, SPORTDiscus *via* EBSCO, Web of Science, PubMed, Scopus, and Cochrane Central Register of Controlled Trials. The search terms were categorized as sample (e.g., adult, athlete, intellectual/developmental disability/impairment, developmental disability); intervention (e.g., RT, fitness, exercise); and testing and assessment protocol (e.g., muscular strength, maximal testing, 1-RM, grip strength, dynamometer, peak torque). A copy of the full search strategy can be found in Supplementary Appendix A. Each database was searched from their earliest available record up to 13 May 2020.

Eligibility criteria

The review included articles that met the following criteria: (1) at least 50% of the participants was diagnosed with an ID (of any

severity) and was between the ages of 18–65 years; (2) the design of the study was a randomized clinical trial (RCT) or a clinical controlled trial (CCT); (3) the design had at least two arms that included an experimental group (EG) and a control group (CG); (4) the EG received a RT intervention and the CG received either no treatment (other than what they typically did during that time) or another type of intervention that did not include RT exercises; (5) the RT intervention included traditional dynamic exercises using coupled concentric and eccentric actions (e.g., free weights, body-weight resistance, elastic tubing, machine weights, isokinetic devices, balance balls); (6) the RT intervention took place over at least two weeks with a minimum frequency of two days per week in order to see a physiological strength change, rather than a neurological improvement in muscle fiber recruitment [37,38]; and (7) the primary outcomes were quantitative measures of maximal static/isometric or dynamic muscular strength. Cochrane's definitions [39] and criteria were used to determine whether a study was an RCT (i.e., the authors stated explicitly that the groups were established by a random allocation) or a CCT (i.e., randomization could not be ruled out or the method of allocation was not considered strictly random). Classification of a study as an RCT or a CCT was based solely on the information reported in the study, not on our inferences; thus, it was not meant to reflect an assessment of the true nature or quality of the allocation procedure. Without the explicit statement of randomization, the study was classified as a CCT.

The exclusion criteria were: (1) more than 50% of participants were reported to have autism, schizophrenia, other psychiatric disorders, acquired cognitive or neurological impairments, sensory impairments, or autoimmune conditions or diseases; (2) any of the participants were children, pregnant women, or individuals with an unmanaged chronic disease or a pathology (including injuries); (3) the CG was exposed to RT exercises during the study (e.g., comparing two different types of RT without a CG); (4) more than 70% of the intervention consisted of aerobic or other non-RT exercises; (5) the RT intervention did not evaluate regimens characteristic of circuit-based RT interventions (e.g., successive sets of different exercises, little rest time, unconventional equipment); (6) the primary outcomes were limited to quantitative measures of submaximal static/isometric or dynamic muscular strength; (7) the article reported only a literature review, qualitative data, or case reports; (8) a full text of the study in English was not available; and (9) the data were from reviews, conference proceedings and abstracts, editorials, dissertations, theses, and articles published in non-peer-reviewed journals.

Study selection

The database searches of article titles yielded 1996 results that were exported to Endnote X9.2 [40]. After duplicates were removed (646 *via* Endnote and 156 manually), the total number of article titles was reduced to 1194, which were assessed for eligibility by two independent authors (CJF and IO). An article title was excluded when two authors (CJF and IO) independently agreed that it failed to meet all inclusion criteria or met an exclusion criterion. The remaining 217 article abstracts were assessed independently for eligibility by two authors (CJF and IO) using the same inclusion/exclusion criteria. If at least one author suggested an abstract should be included, the article's full text was obtained and independently assessed for inclusion by both authors. Disagreements were resolved through discussion or third-party adjudication (RRS) until consensus was achieved. After excluding 171 full-text articles, the reference lists of the remaining 39

articles were screened by one author (CJF) for any additional articles relevant to the topic, as described by Greenhalgh and Peacock [41], which added five articles. Of the 44 articles identified for possible inclusion, the consensus among the four authors excluded 33 articles. Thus, a total of 11 articles remained for final analysis. See Figure 1 for a flow diagram of the search process and reasons for exclusion.

Data extraction and management

Using a pre-specified, pilot-tested form, two authors (CJF and IO) independently coded the following variables in the selected articles: (1) study design (i.e., RCT or CCT); (2) sample (i.e., sex, mean age, and diagnosis); (3) training regimen (i.e., intervention type and setting; exercise type; dose, session duration, and frequency; and sets, reps, and intensity); (4) assessment of strength (i.e., type and regions); (5) familiarization (i.e., number of sessions, type, and strategies used); and (6) findings (i.e., between-group differences at posttest). Coding was cross-checked between coders, and any discrepancies were resolved by mutual consensus or by a third author (RRS). Missing data items were sought by e-mail communication with the corresponding study author. To assess potential coder drift, 30% of the articles were randomly selected for recoding, as described by Cooper et al. [42]. Per-case agreement between two authors (CJF and IO) was determined by dividing the number of variables that were coded the same by the total number of variables. Acceptance required at least 90% agreement. Table 1 provides descriptive characteristics of the 11 analyzed articles.

Risk of bias in individual studies

Two authors (IO and RRS) assessed the risk of bias in the included studies using the Tool for the assEssment of Study qualiTy and reporting in EXercise (TESTEX), which was designed for use by exercise specialists to assess the quality and reporting of exercise training trials [43]. The TESTEX tool uses 12 criteria, with some criteria scoring more than one possible point, for a maximum score of 15 points (5 points for study quality and 10 points for reporting). Consistent with the TESTEX validation study [43], for each article, authors assigned either a value of 0 (absent or inadequately described) or 1 (present or explicitly described) to each TESTEX review item. Both authors had experience in conducting exercise training studies and expertise assessing the study quality of exercise intervention trials. Each author was provided with a copy of the TESTEX protocol, 11 full-text articles, and an Excel spreadsheet to record the data. The inter-observer agreement between the authors was assessed for each individual point available on the TESTEX scale (15 in total) using Cohens Kappa statistic (*K*) [44] and for the total TESTEX score using the intra-class correlation coefficient (ICC). The Kappa's statistic equation is: $K = (P_o - P_c)/(1 - P_c)$, in which P_o is the number of observed agreements and P_c is the number of agreements expected by chance.

Assessment of heterogeneity

As outlined in the Cochrane handbook [39], two authors (IO and ARC) assessed heterogeneity using forest plots, chi-squared (χ^2) tests, and the I^2 statistics. The equation for the I^2 statistics was calculated as $I^2 = 100\% \times (Q - df)/Q$. In this equation, Q is the χ^2 statistic and df is its degrees of freedom. Values of I^2 are percentages of variability in the effect estimates that are due to heterogeneity rather than sampling error. Where evidence of
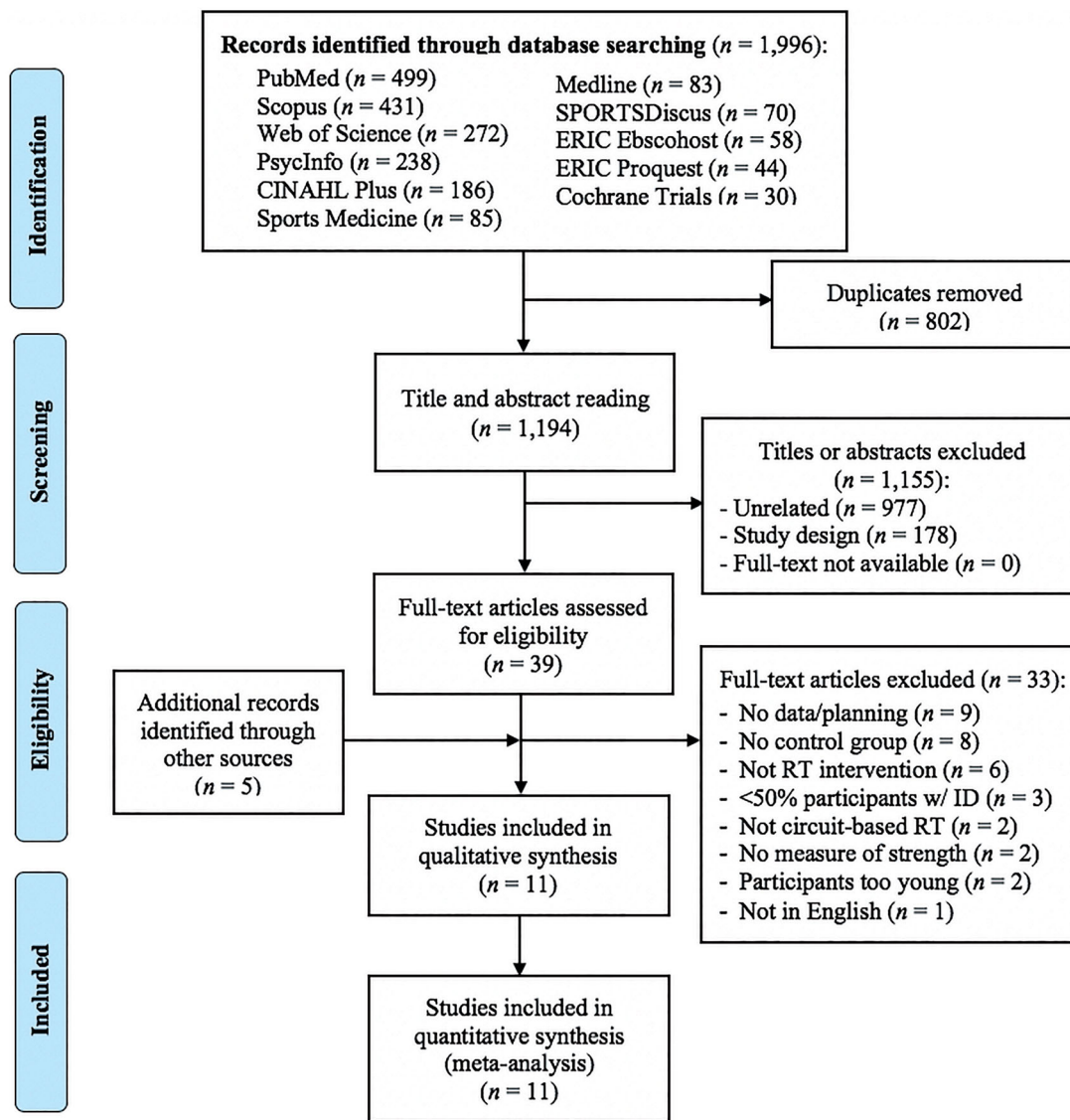
Figure 1. PRISMA flow diagram of included articles in the review and meta-analysis.

substantial ($I^2 > 50\%$) or statistically significant (χ^2, $p < 0.10$) heterogeneity was observed between studies, the data from the outlier studies were extracted and qualitatively investigated.

Measurement of treatment effect

Meta-analyses were performed using the Cochrane's RevMan software, version 5.3 [45], to determine the effect of the RT interventions on muscular strength in adults with ID. For each study, standardized mean differences (SMDs) and 95% confidence intervals (CIs) were calculated using post-intervention continuous data (i.e., comparisons of final values between the EGs and CGs). Treatment effect estimates (SMDs) were weighted using the inverse-variance method and aggregated using a random-effects model due to variability in experimental factors (e.g., training intensity, mode of RT, diagnosis, severity of ID) across the included studies [46]. Data were transformed into the same units when different units of measurement were used for a given outcome variable between studies (i.e., pounds to kilograms). No cases of missing values were found in the included articles. When multiple modes of muscle strength measurements were reported for the primary outcome, because none of the included studies

performed a reliability analysis, the principle of specificity was applied as suggested in the research literature [21,47]. Consequently, priority was given to a post-intervention test that allowed a person to experience conditions similar to those experienced during the RT intervention, in terms of the structure of the test, the mode of contraction, the velocity of contraction and the load(s) or resistance(s) [21]. In some instances, when tests were equally specific, primary analysis variables were prioritized in the following order: lower over upper extremity, single-joint over multiple-joint strength test (to favor low exercise neuromotor difficulty) [47], knee extension over knee flexion [48].

Forest plots were generated to illustrate study-specific effect sizes (ESs) and their 95% CIs, along with the overall pooled effect. Overall effects were considered significant at $p \leq 0.05$, and trends were declared at $p \leq 0.10$. To identify the presence of highly influential studies that might bias the analysis, a sensitivity analysis was carried out for each model by removing one study at a time and then examining the RT intervention predictor. A study was identified as influential if its removal resulted in a change of the predictor from significant or a trend ($p \leq 0.10$) to nonsignificant ($p > 0.10$), or vice versa, or if its removal caused a large change in the magnitude of the coefficient [49]. Sub-analyses were

Table 1. Descriptive characteristics of the studies that assessed effects of RT interventions on measures of muscular strength.

First author (country)	Design	Sample Total (women), M age	Diagnosis (severity)	Intervention type; Setting	Exercise type	Training regimen Dose; Session duration; Frequency	Sets and reps; Intensity (% of 1RM)	Assessment Type; Region (machine)	Familiarization # of sessions; type; strategies	Findings Between-group differences at posttest
RT vs. Control-without-intervention										
Cowley [23] (USA)	CCT	EG = 19 (10), 29±9 y; CG = 11 (3), 27±7 y	DS (mild ID)	PRT; NR	4 UL, 3 LL	10 wks; NR; 2x/wk	3 sets of 8–10 reps; resistance was progressively ↑ to consistently overload the muscle	Isokinetic +isometric PT – KE, KF (Biodex)	1–2 sessions to explain and perform test tasks; Week 1 to explain protocol and tasks.	EG had significantly greater (a) isokinetic PT of KE, KF and (b) isometric PT at 45° and 75°.
Shields [31] (Australia)	RCT	EG = 9 (2), 25.8±5.4y; CG = 11 (5), 27.6±9.5 y	DS (mild to severe ID)	RT; community gym	3 UL, 3 LL, 1 TR	10 wks; NR; 2x/wk	2–3 sets of 10–12 reps; until muscle fatigue, intensity ↑ when 3 sets of 12 reps of an exercise could be completed	1RM – chest and leg press	NR	EG showed a trend toward greater chest press 1-RM.
Shields [52] (Australia)	RCT	EG = 8 (4), 24.5±7.1y; CG = 8 (4), 27.4±9.5 y	PWS (mild to severe ID)	RT; community gym	3 UL, 3 LL, 1 TR	10 wks; NR; 2x/wk	3 sets of 12 reps; 50–60% or until muscle fatigue, intensity ↑ when 2 sets of 12 reps of an exercise could be completed	1RM – chest and leg press	NR	SMD indicated moderate to large ES in favor of the EG for chest press and leg press 1-RM.
RT vs. Control-with-intervention										
Rimmer [50] (USA)	RCT	EG = 12 (NR), 33.8±8 y; CG = 12 (NR), 30.1±4.6 y	ID (IQ = 40–70)	RT; University	5UL, 2 LL	9 wks; 60 min; 2x/wk	3 sets of 8–10 reps at 30%, 60% and 70%, respectively; intensity ↑ by 10lbs or 2.5–5lbs (dependent on exercise) when >12 reps completed during set 3	1RM – leg extension, leg curl, shoulder abduction, pull over, pec deck, biceps, triceps	2 sessions to explain procedures and the use of equipment	EG had significantly greater shoulder abduction, pull over, pec deck, biceps curl, and triceps extension 1-RM.
Suomi [55] (USA)	RCT	EG = 11 (0), 29.9±6.7y; CG = 11 (0), 30.1±5.2 y	ID (EG's IQ = 56.7±6.8; CG's IQ = 58.6±12.6)	RT; University	NR	12 wks; NR; 3x/wk	1 set of NR reps for 2 wks; additional set added each wk; intensity ↑ when reps/set could be performed in <2 secs/rep; intensity ↓ when reps/set could be performed in >3 secs/rep	Isokinetic PT/TW – KE, hip abduction (MERAC)	1 session to explain testing and protocol.	EG had significantly greater (a) PT/TW of hip abduction and (b) TW of KE.
Suomi [54] (USA)	RCT	EG = 6 (0), 27±6.1y; CG=6 (0), 32.6±5.6 y	ID (EG's IQ = 53.7±10.6; CG's IQ = 60.5±11.0)	RT; University	NR	Phase 1: 12 wks; NR; 3x/wk, NR; Phase 2: 52 wks; NR; 2x/wk	Phase 1: 1 set of NR reps for 2 wks; additional set added each wk; intensity ↑ when reps/set could be performed in < 2 secs/rep; intensity ↓ when reps/set could be performed in > 3 secs/rep Phase 2: 2 sets of 20 reps	Isokinetic PT/TW – KE (MERAC)	NR	EG had significantly greater isokinetic PT/ TW of KE after Phase 1 and 2.
RT with other form of exercise vs. control-without-intervention										
Calders [57] (Belgium)	CCT	EG = 15 (9), 42±7.5y; CG = 15 (9), 43±11.4 y	Fragile X s, fetal alcohol s., PWS, hydrocephalus, PDD, Sotos s, DS (EG's IQ = 56±4.3; CG's IQ = 53±5.3)	Aerobic (40%), RT (40%), functional training (20%); day center	1 UL, 2 LL, 1 TR	20 wks; 70 min; 2x/wk	2 sets of 15 reps at 20% for 10 wks; 2 sets of 12 reps at 15% for 10 wks; 2 sets of 8 reps at 10% for 20 wks	1RM – leg extension, leg curl, biceps, triceps, back, abdominal; Handgrip	NR	EG had significantly greater (a) 1-RM in all muscles and (b) grip strength.
Ko [56] (South Korea)	RCT	EG = 10 (NR), 45.2±2.5y; CG = 10 (NR), 49.8±3 y	ID (NR)	Aerobic, RT; NR	NR	24 wks; 30–60 min; 3–4x/wk	55–69% HRmax; exercise strength was analyzed via cardiac rates during exercise via a Polar heart rate tester	Isokinetic PT – KE, KF (Cybex)	Explained test and intervention procedures and the use of equipment; rehearsed procedures 3x	Between-group differences NR. Only EG had significantly greater isokinetic PT of KE, KF at posttest.

Table 1. Continued.

First author (country)	Design	Sample: Total (women), M age	Sample: Diagnosis (severity)	Training regimen: Intervention type; Setting	Exercise type	Dose; Session duration; Frequency	Sets and reps; Intensity (% of 1RM)	Assessment Type; Region (machine)	Familiarization # of sessions; type; strategies	Findings Between-group differences at posttest
Oviedo [58] (Spain)	CCT	EG = 37 (15), 41±11y; CG = 29 (12), 46±12y	DS, West s., Cornelia Lange s.; Microcephaly; Lennox s.; CP; Conduct disorder; epilepsy; autism (mild to moderate ID)	Aerobic (25–50%), RT (25–35%), Balance (15–25%); NR	3 UL, 3 LL, 1 TR	14 wks; 60 min; 3x/wk	2 sets of 15 reps for 4 wks; 2 sets of 12 reps for 5 wks; 2 sets of 10–12 reps for 5 wks; 1 set until exhaustion for all wks	Foot-plate dynamometer – leg extension; Handgrip	2–3 sessions to explain testing; Week 1–2 taught how to use equipment safely.	EG had significantly greater (a) leg strength and (b) grip strength.
Rimmer [51] (USA)	RCT	EG = 30 (16), 39±6.2y; CG = 22 (13), 40.6±6.5y	DS (NR)	Aerobic (50–75%), RT (25–35%); University	6 UL, 2 LL	12 wks; 45–60 min; 3x/wk	1 set of 10–20 reps at 70%; intensity ↑ by 10% when able to complete 20 reps for 2 consecutive sessions	1RM – chest and leg press; Handgrip	30 min demonstrating correct technique on strength-testing machines followed by performing 1–2 sets of 8–10 reps	EG had significantly greater leg press and bench press 1-RM.
Van Schijndel-Speet [53] (The Netherlands)	RCT – cluster	EG = 66 (38), 58.2 y±NR; CG = 65 (34), 57.9 y±NR	DS, autism, behavioral problems, unknown (mild to severe ID)	Aerobic, RT, Balance; day center	18 exercises (type NR)	32 wks; 15–45 min; 3x/wk	NR	Handgrip	NR	EG had significantly greater grip strength.

RT: resistance training; EG: experimental group; CG: control group; DS: Down syndrome; CP: cerebral palsy; PDD: pervasive developmental disorder; 1-RM: one-repetition maximum; PWS: Prader–Willi syndrome; s.: syndrome; UL: upper limbs; LL: lower limbs; TR: trunk; PT: peak torque; TW: total work; KE: knee extension; KF: knee flexion; SMD: standardized mean difference; NR: not reported.

performed to determine the differences in effects of RT interventions on muscular strength based on the exercise regimen (i.e., RT with and without another form of exercise).

Results

Study and subject characteristics

Eleven studies from published articles met the inclusion criteria for this systematic review and meta-analysis (see Figure 1). The publication dates of the articles ranged from 1991 to 2020. Details of the included articles are provided in Table 1. Five studies were conducted in the United States, two in Australia, and one each in Belgium, South Korea, Spain, and the Netherlands. Of the 11 studies, eight were reported by the authors as RCTs [31,50–56] and three as CCTs [23,57,58].

Group sample sizes in the selected studies ranged from 6 to 53 participants ($M = 17.6$; $SD = 12.9$). Collectively, the analysis included 388 participants, 210 in the EGs and 178 in the CGs, which indicates a large disparity between the number of participants in the groups. The mean age was similar between the EGs ($M = 35.9$, $SD = 10.3$, range = 24.5–58.2) and the CGs ($M = 37.5$, $SD = 10.6$, range = 27.0–57.9). Of the 388 participants, 174 (45%) were women and 214 (55%) men. While two studies [54,55] collected data from men only, the remaining studies included both women and men. Three studies [23,31,51] collected data exclusively from adults with Down syndrome and one [52] exclusively from adults with Prader–Willi syndrome. Four studies [50,54,55,57] reported IQ scores for each group. Only Suomi's studies [54,55] identified the cognitive test (i.e., Stanford-Binet Intelligence Quotient Test) that was used to assess IQ. Three studies [53,57,58] collected data from a mixed sample of different developmental disabilities (see Table 1) that the researchers reported to range in ID severity from mild to moderate [57,58] or mild to severe [53]. One study [56] involved adults with ID, but the researchers omitted detail about the participants' diagnoses.

RT protocol characteristics

Six studies evaluated the effects of an intervention comprised only of RT [23,31,50,52,54,55], three of which utilized an active CG that participated either in a university-based program consisting of dance (35 min), aquatics (35 min), and lifetime sports (35 min) [50] or a day center-based program consisting of non-strength training activities such as card playing, dart throwing, shuffleboard playing, and horse-shoe throwing with the researchers [54,55]. Five studies evaluated the effects of an intervention that combined RT with other forms of exercise, such as aerobic training [51,53,56–58], balance training [53,58], or functional training [57]. CG participants in the eight studies that did not utilize an active-CG intervention were asked to engage in normal daily activities. RT interventions were most commonly delivered in a university-based setting [50,51,54,55], followed by a community-based gym [31,52] and a day center for adults with ID [53,57]. Three articles did not report the intervention setting.

The exercise regimens ranged widely across the RT interventions in the 11 studies. All articles reported the training dose and frequency of sessions. The training dose ranged from 9 to 64 weeks ($M = 19.7$, $SD = 16.4$). Training frequency ranged from 2 to 4 sessions per week ($M = 2.6$, $SD = 0.7$). Session duration was reported in only six articles [50,51,53,56–58], with a median session duration between 30 and 70 min (including warm-up and cool-down) ($M = 52.9$, $SD = 14.0$). Two articles [53,56] did not provide any detail on the intensity of exercise during the RT

intervention. Of the nine articles that provided some detail about exercise intensity, only four of them [50–52,57] reported the percentage of 1-RM for intensity, which ranged between 10% and 70%. Seven articles [23,31,50–52,57,58] reported the types of exercise that were utilized in their RT interventions. All seven RT interventions included upper-limb and lower-limb muscle-group exercises and three [52,57,58] additionally included abdominal muscle-group exercises. The upper-limb exercises were chest press, pectoral deck, seated row, lat pull-down/over, shoulder press and abduction, lateral deltoids, biceps curls, and triceps extensions. The lower-limb muscle-group exercises were leg press, leg extension, leg curls, seated row, and calf raise. When all exercises from the seven studies were combined, single-joint ($n = 22$) and multiple-joint ($n = 24$) exercises were equally represented. RT exercises were performed using a variety of equipment, including traditional fitness-center weight machines [23,31,50–52,57] and hydraulic RT equipment/machines [54,55]. Exercises in Oviedo et al. [58] were performed with free weights, elastic bands, and medicine balls.

Only six articles [23,50,51,55,56,58] reported, either briefly or extensively, the procedures that the researchers used to familiarize participants with the muscular-strength-test tasks, the equipment, and the study protocols. The familiarization phases lasted either 1–2 sessions [23,50,51,55] or 2–3 sessions [56,58] for an average of 1.9 sessions ($SD = 0.7$).

Outcome measures

Isotonic muscular strength was measured with 1-RM tests and dynamometer-based tests. Chest-press and leg-press 1-RM scores were reported in three articles [31,51,52] and leg curl, leg extension, biceps curl, and triceps extension 1-RM scores were reported in two articles [50,57]. Knee-extension foot-plate dynamometer scores were reported in one study [58]. Isokinetic muscular strength was measured with Biodex, Cybex, or MERAC dynamometers. Knee-extension and -flexion peak torque or total work scores were reported in four studies [23,54–56]. Isometric muscular strength was measured with dynamometer-based tests. Handgrip dynamometer scores were reported in four studies [51,53,57,58] and knee-extension and -flexion peak torque Biodex dynamometer scores were reported in one study [23].

Reliability analyses were performed in three studies [50,54,55]. Of those, the two Suomi et al. articles [54,55] reported high intra-class reliability coefficients ($r = 0.82–0.97$) for the pretest and posttest isokinetic measures of peak torque and total work of knee extension or hip abduction, and Rimmer et al. [50] reported high test-retest reliability coefficients ($r = 0.95$) for the leg-curl 1-RM.

Quality assessment

A detailed description of the TESTEX scores obtained independently by two authors (IO and RRS) is shown in Table 2. There were 13 coding discrepancies between these two authors, which were reviewed and classified as factual ($n = 5$) or interpretative ($n = 8$) errors. Factual errors were regarded as transcription errors where the correct answer was available in the article and either not identified by the authors or inaccurately reported by the researchers. Factual errors were corrected by both authors after revisiting the articles. Interpretative errors were regarded as errors in which study information was inferred, unclear, or the definition of the coding criterion was perceived to be ambiguous, and the authors had to interpret the information. Interpretative errors

Table 2. Assessment of quality using TESTEX scale of included articles in meta-analysis (ordered according to the overall TESTEX score).

Study (publication year)	Eligibility criteria specified	Method of randomization specified	Allocation concealed	Groups similar at baseline	Blinding of assessors	Outcome measures in >85% participants[a]	Intention-to-treat analysis	Between-group statistical comparisons reported[b]	Point measures and measures of variability reported	Activity monitoring in control group	Relative exercise Intensity reviewed	Exercise volume and energy expended	Overall TESTEX Score (out of 1/15)
Shields (2008)	1	1	1	1	1	3	1	2	1	0	1	1	14
Shields (2020)	1	1	1	1	1	3	1	2	1	0	1	0	13
Calders (2011)	1	0	0	1	1	3	1[c]	2	1	0	1	1	12
Van Schijndel-Speet (2017)	1	1	1	1	0	2	0	2	1	0	0	1	10
Oviedo (2014)	1	0	0	0	0	2	0	2	1	0	1	1	8
Rimmer (2004)	1	0	1	1	0	0	0	2	1	0	1	0	7
Suomi (1998)	1	0	0	1	0	2	0[c]	1	1	0	1	0	7
Suomi (1995)	1	0	0	0	0	1	1[c]	2	1	0	1	0	7
Rimmer (1991)	0	0	0	0	0	2	1[c]	1	1	0	1	0	6
Cowley (2011)	1	0	0	0	0	0	0	2	1	0	1	0	5
Ko (2012)	0	0	0	1	0	0	0	0	1	0	0	0	2

[a]Three points possible; [b]Two points possible; [c]If intention-to-treat was not specifically mentioned, but it was noted that no participants withdrew and all data were analyzed, 1 point was awarded; Grey criteria evaluated study quality (out of 5 points); White criteria evaluated study reporting (out of 10 points).
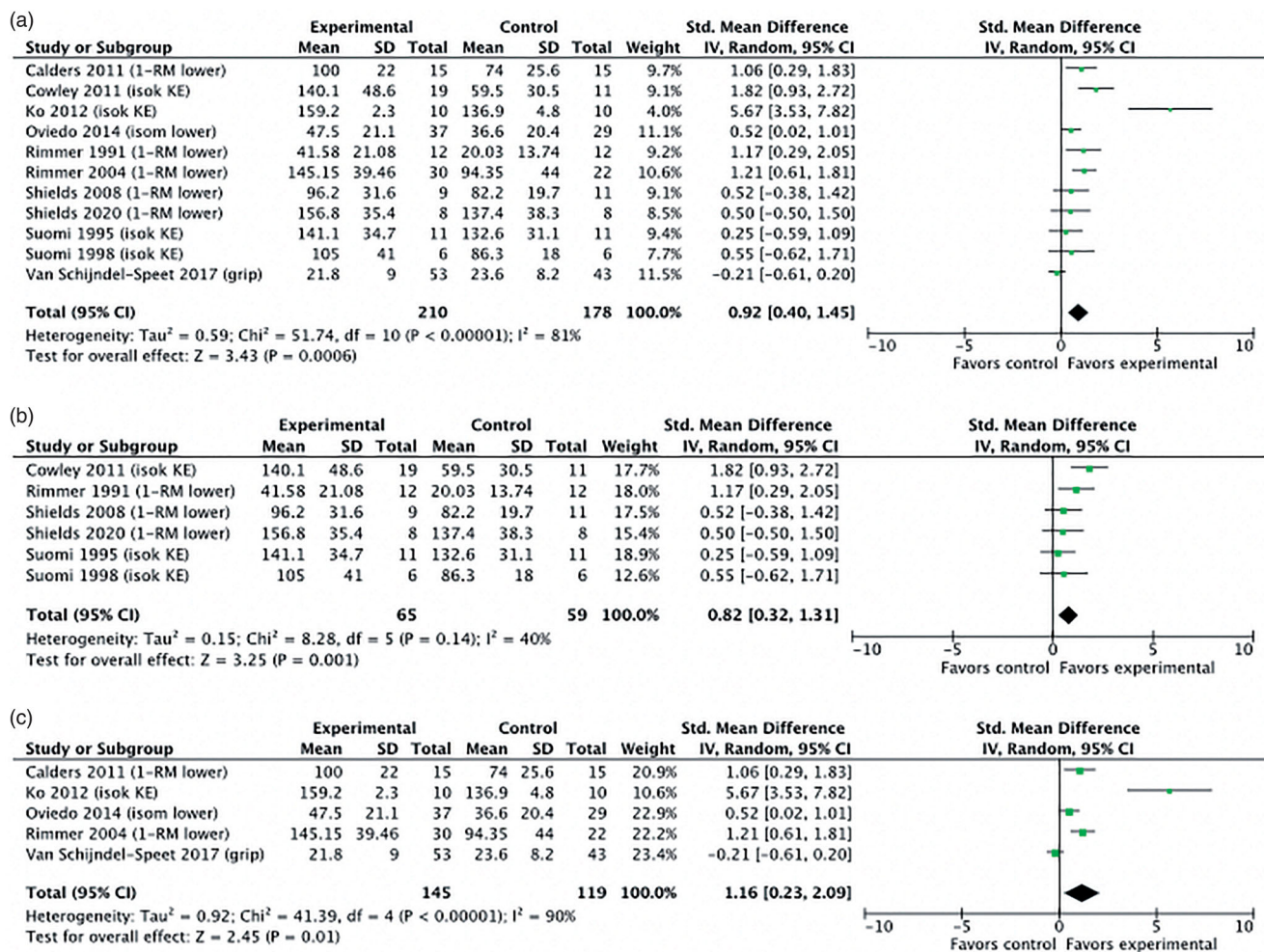
(a)

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Calders 2011 (1-RM lower)	100	22	15	74	25.6	15	9.7%	1.06 [0.29, 1.83]
Cowley 2011 (isok KE)	140.1	48.6	19	59.5	30.5	11	9.1%	1.82 [0.93, 2.72]
Ko 2012 (isok KE)	159.2	2.3	10	136.9	4.8	10	4.0%	5.67 [3.53, 7.82]
Oviedo 2014 (isom lower)	47.5	21.1	37	36.6	20.4	29	11.1%	0.52 [0.02, 1.01]
Rimmer 1991 (1-RM lower)	41.58	21.08	12	20.03	13.74	12	9.2%	1.17 [0.29, 2.05]
Rimmer 2004 (1-RM lower)	145.15	39.46	30	94.35	44	22	10.6%	1.21 [0.61, 1.81]
Shields 2008 (1-RM lower)	96.2	31.6	9	82.2	19.7	11	9.1%	0.52 [-0.38, 1.42]
Shields 2020 (1-RM lower)	156.8	35.4	8	137.4	38.3	8	8.5%	0.50 [-0.50, 1.50]
Suomi 1995 (isok KE)	141.1	34.7	11	132.6	31.1	11	9.4%	0.25 [-0.59, 1.09]
Suomi 1998 (isok KE)	105	41	6	86.3	18	6	7.7%	0.55 [-0.62, 1.71]
Van Schijndel–Speet 2017 (grip)	21.8	9	53	23.6	8.2	43	11.5%	-0.21 [-0.61, 0.20]
Total (95% CI)			**210**			**178**	**100.0%**	**0.92 [0.40, 1.45]**

Heterogeneity: Tau² = 0.59; Chi² = 51.74, df = 10 (P < 0.00001); I² = 81%
Test for overall effect: Z = 3.43 (P = 0.0006)

(b)

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Cowley 2011 (isok KE)	140.1	48.6	19	59.5	30.5	11	17.7%	1.82 [0.93, 2.72]
Rimmer 1991 (1-RM lower)	41.58	21.08	12	20.03	13.74	12	18.0%	1.17 [0.29, 2.05]
Shields 2008 (1-RM lower)	96.2	31.6	9	82.2	19.7	11	17.5%	0.52 [-0.38, 1.42]
Shields 2020 (1-RM lower)	156.8	35.4	8	137.4	38.3	8	15.4%	0.50 [-0.50, 1.50]
Suomi 1995 (isok KE)	141.1	34.7	11	132.6	31.1	11	18.9%	0.25 [-0.59, 1.09]
Suomi 1998 (isok KE)	105	41	6	86.3	18	6	12.6%	0.55 [-0.62, 1.71]
Total (95% CI)			**65**			**59**	**100.0%**	**0.82 [0.32, 1.31]**

Heterogeneity: Tau² = 0.15; Chi² = 8.28, df = 5 (P = 0.14); I² = 40%
Test for overall effect: Z = 3.25 (P = 0.001)

(c)

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Calders 2011 (1-RM lower)	100	22	15	74	25.6	15	20.9%	1.06 [0.29, 1.83]
Ko 2012 (isok KE)	159.2	2.3	10	136.9	4.8	10	10.6%	5.67 [3.53, 7.82]
Oviedo 2014 (isom lower)	47.5	21.1	37	36.6	20.4	29	22.9%	0.52 [0.02, 1.01]
Rimmer 2004 (1-RM lower)	145.15	39.46	30	94.35	44	22	22.2%	1.21 [0.61, 1.81]
Van Schijndel–Speet 2017 (grip)	21.8	9	53	23.6	8.2	43	23.4%	-0.21 [-0.61, 0.20]
Total (95% CI)			**145**			**119**	**100.0%**	**1.16 [0.23, 2.09]**

Heterogeneity: Tau² = 0.92; Chi² = 41.39, df = 4 (P < 0.00001); I² = 90%
Test for overall effect: Z = 2.45 (P = 0.01)

Figure 2. Forest plot of the results of a random effects meta-analysis shown as standardized mean differences (SMD) with 95% confidence interval (CI) for (a) all the studies combined, (b) the non-combined-intervention studies that exclusively utilized RT exercises, and (c) the combined-intervention studies that utilized RT with other forms of exercise. The shaded circle represents the point estimate for each individual study and the weight of the study in the meta-analysis. The diamond represents the overall mean difference of the studies. Grip: grip strength; Isok: isokinetic strength; Isom: isometric strength.

were discussed with another author (ARC) and the appropriate code was determined by a simple majority decision. The agreement (K) between the two authors (IO and RRS) ranged between 0.48 (moderate) and 1.00 (perfect). Moderate agreements occurred in two categories: "Method of randomization specified" and "Between-group statistical analysis reported for secondary outcomes." Substantial agreements occurred in four categories: "Adherence >85%," "Relative exercise intensity remained constant," "Blinding of assessor," and "Allocation concealment." Almost perfect or perfect agreement occurred in the remaining nine categories. The correlation between the two authors on the overall TESTEX scores was statistically significant (ICC = 0.97, 95% CI = 0.88–0.99, $p < 0.000$).

The total TESTEX score (out of 15), study quality subscore (out of 5), and study reporting subscore (out of 10) of the 11 studies were 8.3 ± 3.6 (range 2–14, $Mdn = 7$), 2.5 ± 1.7 (range 0–5, $Mdn = 2$), and 5.8 ± 2.4 (range 1–9, $Mdn = 6$), respectively. The most common study quality concerns were the lack of reporting on the method of randomization (73% of studies), the lack of blinding of assessors (73%), and the lack of allocation concealment (64%). The most common study reporting concerns were the lack of activity monitoring in the CG (studies), the lack of intention-to-treat (73%), the lack of reporting on adverse events

(64%), and the lack of reporting on exercise volume and energy expenditure (64%).

Effect of RT interventions

Table 1 provides an overview of the between-group effects at post-intervention as they were reported in the articles in this review. Of the 11 articles, eight [23,50,51,53–55,57,58] reported statistically significant group differences and two [31,52] reported trends toward significant group differences in muscular strength at post-intervention in favor of the EG. While one article [56] reported that only the EG had significantly greater isokinetic peak torque gains at posttest, it does not appear that the authors performed any post-intervention between-group analyses. This article also had the lowest overall TESTEX quality score (i.e., 2 out of 15).

The effects of RT interventions implemented in the 11 studies were evaluated in a series of meta-analyses that compared post-intervention strength scores between participants in the EGs and the CGs (see Figure 2). Results of the meta-analysis in Figure 2(a) show a statistically significant effect of the 11 RT interventions on muscular strength of adults with ID (SMD = 0.92; 95% CI = 0.40 to 1.45; $z = 3.43$; $p = 0.000$). This significant effect was substantially

heterogeneous ($I^2 = 81\%$, $p < 0.000$), but sensitivity analyses did not reveal any influential studies.

The next analysis explored the effects of the two types of RT interventions (i.e., combined vs. non-combined). Figure 2(b,c) provide results of meta-analyses performed separately for the six combined and five non-combined RT interventions, respectively. The results revealed that both sets of studies produced statistically significant gains in muscular strength, with the non-combined RT interventions being more effective (SMD = 0.82; 95% CI = 0.32 to 1.31; $z = 3.25$; $p = 0.001$) compared with the combined interventions (SMD = 1.16; 95% CI = 0.23 to 2.09; $z = 2.45$; $p = 0.01$). Interestingly, while the non-combined-intervention studies had non-significant heterogeneity ($I^2 = 40\%$, $p = 0.140$), the combined-intervention studies had considerable heterogeneity ($I^2 = 90\%$, $p < 0.000$). Sensitivity analysis revealed an influential combined-intervention study [56], which also had the lowest TESTEX quality score (2 out of 15). Removal of the study slightly reduced the heterogeneity ($I^2 = 84\%$, $p = 0.000$) and changed the magnitude of the significant difference between the EG and CG (SMD = 0.61; 95% CI = −0.07 to 1.29; $z = 1.75$; $p = 0.08$). There was an insufficient number of studies to quantify the magnitude of effect based on the modality of testing.

Discussion

The current systematic review with a meta-analysis evaluated the effects of RT interventions on muscular strength in adults with ID. A total of 11 clinical trials met the inclusion criteria. The principal finding was that the RT interventions had an overall significant ($p \leq 0.05$) and large effect on muscular strength in adults with ID, ages 25–58 years. This demonstrates that this training modality increased muscular strength independent of the particular research design or the intervention protocol used in the studies. Therefore, the hypothesis that RT interventions will have significant effects ($p \leq 0.05$) on measures of muscular strength in adults with ID was supported. However, even though two studies [31,52] had almost perfect TESTEX scores, the overall moderate quality of the synthesized studies with almost 50% of the studies scoring below 50% of the maximum possible TESTEX score, the small sample sizes, and the considerable heterogeneity between the studies indicate that there is a potential risk of bias. Therefore, caution should be taken in the interpretation of the meta-analysis results.

A positive finding of this review is that a majority of studies (8 of 11) were RCTs (as reported by their authors), which is consistent with the findings of another review of exercise interventions for adults with ID [28]. A less positive finding is that only three out of the eight RCTs specified the methods/techniques used to randomize participants into groups or collected data by blinded assessors, and only four concealed allocation. Similar design flaws were found in a review by Jeng, Chang, Liu, Hou, and Lin [59] that evaluated the effects of broad exercise interventions on skill-related physical fitness in adolescents with ID. Other design flaws were related to small or unequal sample sizes. Two CCTs [23,58] and one RCT [51] included a larger number of participants in their EGs than in their CGs. While Cowley et al. [23] attributed unequal sample-sizes to participant recruitment (EG participants were recruited from two sites and CG participants from one site based on the sites' capabilities to provide a supervised intervention), Oviedo et al. [58] attributed it to an attrition of participants in the CG (i.e., three participants dropped out because they changed their workplace and three refused to be part of the inactive CG). Rimmer et al. [51] did not provide details about group allocation

to determine whether the use of simple rather than complete randomization contributed to the unequal sample sizes. However, the combination of power analysis carried out in two of the studies [23,51] and the overall larger sample sizes in the three studies minimize the probability of Type II error in this review [60]. To ensure the groups are comparable, future research reports should precisely specify the method of randomization they followed to allocate participants to groups (e.g., computer-generated vs. hat-drawn, complete vs. simple) and attempt to conceal allocation so that neither the participants nor the caregivers or the researchers are aware to which group a person is being allocated prior to consenting to participate in the study [43]. 0Although blinding of participants and researchers is very difficult to implement in exercise training studies, it is reasonable to expect researchers to blind assessors who conduct outcome data measurements [43].

Other reporting concerns were the lack of activity monitoring in the CG, the lack of intention-to-treat, and the lack of reporting adverse events, exercise volume, and energy expenditure. It has been recommended that robust study designs should monitor and then compare PA levels between the EG and CG to capture inadvertent changes in PA and avoid crossover to exercise [43]. This could entail providing caregivers or the participants a diary to log their PA or a more precise, objective assessment method such as accelerometers or pedometers [61]. However, more research is needed with adults with ID to identify optimal approaches to assessing PA, particularly in a CG given the potential for measurement reactivity (measurement causes alterations in those being measured), a fairly common bias in PA studies [62,63]. Further, due to their poor compliance [61], there is a need to establish and standardize specific accelerometer/pedometer protocols for measuring PA levels in persons with ID for higher quality and more comparable data [64].

The RT interventions evaluated in this review were predominantly of a short-term duration, with seven ranging from 9 to 14 weeks ($M = 11.0$, $SD = 1.7$) and five [54,56–58,65] ranging from 20 to 64 ($M = 35.0$, $SD = 20.0$). While studies have shown muscle hypertrophy and improved dynamic muscular strength can occur after six weeks of high-intensity RT in novice participants [38], the short-term nature of the RT interventions makes it difficult to draw conclusions regarding their long-term effects on muscular strength or other health- or skill-related outcomes. Similar to the findings of Bouzas et al. [28], the most common interventions involved the performance of circuit-based progressive RT exercises using weight-lifting machines. "Circuit-based" training refers to a number of carefully selected exercises arranged in a specific successive order [66]. The use of weight-lifting machines in RT is particularly useful for persons with ID because they promote faster acquisition and more consistent performance of RT exercises [67–69]. Other equipment used in the RT circuits of the included studies were free weights, elastic bands, or medicine balls. Only two studies provided information on how they adapted their RT intervention protocol to accommodate the diverse needs of their participants with IDs. For example, in Rimmer et al. [50], higher-functioning participants were paired with lower-functioning participants to help them record scores. Additionally, they labeled machines by the movement for easier recognition. In Suomi et al. studies [54,55], each participant selected an activity in which they participated in each training session. Only van Schijndel-Speet et al. [53] systematically developed their PA program consisting of an RT component using an intervention mapping approach, which involved interviewing care providers, exercise experts, and older adults with mild to moderate ID.

Given the considerable heterogeneity among the 11 studies, subgroup analyses were performed to determine whether the two types of intervention (i.e., combined vs. non-combined RT interventions) had different effects on muscular strength in adults with ID. In 45% of the studies, RT exercises were combined with balance exercises (e.g., balance boards/pads) and/or aerobic exercises (e.g., stationary biking, brisk walking, jogging, running, and aerobic dancing). The selection of aerobic exercises was limited in some studies based on the abilities of the participants. For example, Rimmer et al. [51] noted that their older participants had difficulty with or refused to walk on a treadmill and instead chose to use a stationary bicycle. Balance plays an important role in mobility as well as stability [70]. With advanced age, muscular strength around the knees and ankles decreases and postural control mechanism becomes less efficient, resulting in gait and balance impairments and increased risk of falls [70]. The choice of a stationary bicycle over a treadmill among some older participants might have reflected the higher prevalence of falls in this population group [71]. Both combined and non-combined RT interventions produced improvements in muscular strength. However, sensitivity analysis showed that removing the study by Ku et al. [56] from the combined-interventions category substantially altered the magnitude of the differences between conditions to non-significant. The relatively low number of studies limited statistical power to draw firm inferences, but examination of the adjusted CI (−0.07 to 1.29) indicates a likely benefit in favor of non-combined RT interventions. It should be noted, however, that the four included studies in the analysis had potential confounding variables (e.g., a longer length of the combined RT interventions) that may have impacted results. The extent and direction to which these factors may have influenced improvements in muscular strength is unclear and warrants further investigation.

Although the findings of this review revealed positive effects, participant characteristics varied considerably across studies, which also could have contributed to the considerable heterogeneity. For example, a larger proportion of the samples in some of the studies were male participants, in two cases [54,55] to the complete exclusion of female participants. Unequal gender representation is consistent with the higher prevalence of ID in male vs. female persons [72]. Further, participants had a wide range of diagnoses with some studies exclusively targeting a specific genetic phenotype such as Down syndrome or Prader–Willi syndrome. Those conditions have specific biological and behavioral symptoms in addition to ID that might have contributed to the heterogeneity and limitations of generalizability to all adults with ID. Similarly, participants had a wide range of ID severity (mild to severe), which were verified with a psychological test in only two studies [54,55]. The ID definition specifies at least three required elements to diagnose a person with ID—IQ, adaptive behavior, and age of onset [73]; however, such information was not reported in most articles. Additionally, most of the participants were volunteers and in two of the four CCTs either the care organization [23] or the participant [58] chose to be allocated to the EG or the CG. These factors could have contributed to selection bias toward adults with ID with less pervasive support needs and higher motivation to participate. To minimize selection bias, researchers should make their study representative by recruiting larger sample sizes and by closely matching participants based on their support needs in the groups. The multidimensional nature of IDs necessitates that researchers assess (and report) each of the three elements of ID using psychometrically sound and individually administered instruments [74] and collect information on the person's motivation to participate in the RT intervention. If exclusionary criteria are used, they need to be carefully justified and considered when making conclusions.

Accurate measurement of muscular strength is essential for the development and evaluation of effective RT interventions that improve muscular strength [75]. The most commonly used measures of maximum muscular strength in the included studies were 1-RM tests (mostly chest press and leg press), followed by isokinetic tests (mostly leg-extension peak torque), and then isometric tests (mostly handgrip), which is consistent with measures used in RT intervention clinical trials with general populations [76]. A recent review [75] concluded that the 1-RM test is a reliable indicator of isotonic muscular strength in general populations, "regardless of resistance training experience, number of familiarization sessions, exercise selection, part of the body assessed (upper vs. lower body), and sex or age of participants" (p. 14). Likewise, isokinetic and isometric dynamometer tests have been shown to be highly reliable and internally valid indicators of isokinetic and isometric muscular strength in general populations [47,77]. However, evidence of the reliability of these three sets of tests in clinical populations is scarce [75]. Although feasibility and test-retest reliability was found to be high for the isokinetic peak-torque test and the handgrip test in a limited number of studies with adults with ID [e.g., 78], 1-RM test reliability analyses were not performed in other studies with this population [75] nor in the five studies included in this review. Additionally, there are conflicting findings in the literature about whether or not isometric testing is predictive of dynamic performance [21]. Because we know this population is heterogeneous but we do not know how the authors of the included studies modified their testing protocols to accommodate their participants' relevant differences, it is reasonable to believe that the validity of their muscular strength tests could have been compromised. Clinicians and researchers should recognize potential sources of testing error when measuring muscular strength to evaluate interventions.

Growing evidence suggests that muscular strength testing can be influenced substantially by variables such as a participant's motivation, performance of the RT tasks, and capacity to tolerate maximal loads [79–81]. None of the included studies, however, reported a detailed testing protocol to permit us to determine whether those important variables were addressed. Further, only two studies for the 1-RM tests [50,51] and four studies for the isokinetic tests [23,55,56,58] reported to have provided 1–2 familiarization sessions, which were mostly limited to instruction on how to use equipment prior to the testing with no rehearsal by participants or corrective feedback by researchers. No familiarization was reported in studies that collected data with isometric tests. Correct performance of RT exercises is paramount for RT interventions to be safe and effective and for the evaluation of their efficacy to be accurate [79,80,82]. Due to their limitations in intellectual functioning and adaptive skills [73], the need for familiarization is particularly important for adults with ID [67,68,83]. With ample support from the evidence base, we encourage researchers and clinicians to implement adequate familiarization training prior to muscle strength testing or an RT intervention/program. Simple and clear verbal directions should be provided along with demonstrations of the exercises that the participants are being asked to perform [69]. The use of systematic prompting and positive reinforcement will promote their accurate performance of the RT exercise tasks [67,68], as well as their ability to apply effort when working against resistance. Modifications of the tasks or equipment might be needed to accommodate the needs of some participants [67,69]. Exercise intensity can be prescribed and monitored with a rating of perceived exertion scale that uses

modified and condensed verbal anchors with images reflecting the various intensity categories at each level of exertion [69]. In addition, clinicians should include adequate time to thoroughly explain the exertion scale, which may be needed to ensure that a participant understands exactly what is being asked. By including these individualized exercise supports in RT interventions, clinicians can greatly improve the ability of persons with ID to independently manage their own exercise behavior with the goal of achieving healthy levels of physical activity and physical fitness [67]. Ultimately, these lifestyle changes have the potential to provide meaningful benefits across the recreational and vocational areas.

To our knowledge this is the first systematic review with a meta-analysis of published, peer-reviewed clinical trials that evaluated the effects of RT interventions on muscular strength in adults with ID. The criteria for inclusion were determined and published at PROSPERO prior to the search to minimize reviewer bias. This objectivity was strengthened by adherence to the PRISMA reporting guidelines. Three limitations need to be noted. First, only studies that used measures of maximum muscular strength were included. To achieve a more comprehensive picture, future reviews should consider studies that explore the effects of RT interventions on measures of muscular endurance, particularly since evidence from research with general populations supports the interference of endurance training exercises on RT-induced muscle hypertrophy and strength [84]. Second, at this time, the number of existing studies is inadequate to perform subgroup meta-analyses that would compare the effects based on the different demographic, intervention, or outcome measure variables. This limits generalizability to all adults with ID as well as formulation of exercise guidelines for specific populations. Third, there may be publication bias in the selection of studies as abstracts, or studies published in non-peer-reviewed or non-English-written journals were excluded.

Conclusions

This systematic review and meta-analysis found that RT interventions can greatly increase muscular strength in adults with ID, independent of the protocols used in the included studies. Interventions that combined RT exercises with aerobic or balance exercises demonstrated higher heterogeneity and lower effect compared to non-combined RT interventions. Even though the results of this review support the hypothesis that RT interventions will have significant effects ($p \leq 0.05$) on measures of muscular strength in adults with ID, the considerable heterogeneity of the effects, the small number of RT interventions (most of which were of short duration and with small sample sizes), the poorly described or executed familiarization, intervention, and measurement protocols, and the compromised methodological quality of the eligible studies warrant additional investigations. Careful consideration should always be given to the individual needs of persons with ID when prescribing a fitness regimen for them. Due to their limitations in intellectual functioning and adaptive skills [73], adults with ID may need individualized support in the form of adaptations and extended practice to learn how to perform the exercises correctly and become familiar with the testing or RT regimen procedures [68,85,86]. Nevertheless, this review provides valuable information to guide researchers and clinicians at this point in time in the design of exercise RT interventions that effectively promote muscular strength in adults with ID and thereby facilitate their realization of the range of associated health benefits. Further intervention research is needed that

adequately reports randomization and concealment allocation procedures, uses blind assessors, reports attendance and adverse events, tracks energy expenditure, monitors activity of all participants, and assesses the long-term effects of RT interventions on muscular strength.

ORCID

Iva Obrusnikova http://orcid.org/0000-0001-6098-453X
Cora J. Firkin http://orcid.org/0000-0001-8053-6443
Richard R. Suminski http://orcid.org/0000-0001-9401-8620

References

[1] Draheim CC. Cardiovascular disease prevalence and risk factors of persons with mental retardation. Ment Retard Dev Disabil Res Rev. 2006;12(1):3–12.

[2] Sutherland G, Couch MA, Iacono T. Health issues for adults with developmental disability. Res Dev Disabil. 200223(6): 422–445.

[3] Dairo YM, Collett J, Dawes H, et al. Physical activity levels in adults with intellectual disabilities: a systematic review. Preventive Med Rep. 2016;4:209–219.

[4] U.S. Department of Health and Human Services. Physical activity guidelines for Americans. In: Promotion OoDPaH, editor. 2nd ed. Washington (DC): U.S. Department of Health and Human Services; 2018.

[5] World Health Organization. Global action plan on physical activity 2018-2030: more active people for a healthier world. Geneva: World Health Organization; 2018.

[6] American College of Sports Medicine. ACSM's health-related physical fitness assessment manual. 5th ed. Philadelphia (PA): Wolters Kluwer; 2018.

[7] Corbin CB, Pangrazi RP, Franks BD. Definitions: health, fitness, and physical activity. Pres Counc Phys Fitness Sports Res Digest. 2000;3(9):1–6.

[8] Graham A, Reid G. Physical fitness of adults with an intellectual disability: a 13-year follow-up study. Res Q Exerc Sport. 2000;71(2):152–161.

[9] Oppewal A, Hilgenkamp TIM, van Wijck R, et al. Cardiorespiratory fitness in individuals with intellectual disabilities-a review. Res Dev Disabil. 2013;34(10):3301–3316.

[10] Croce RV, Pitetti KH, Horvat M, et al. Peak torque, average power, and hamstrings/quadriceps ratios in nondisabled adults and adults with mental retardation. Arch Phys Med Rehabil. 1996;77(4):369–372.

[11] Zghal F, Borji R, Colson SS, et al. Neuromuscular characteristics in trained vs. sedentary male adults with intellectual disability. J Intellect Disabil Res. 2019;63(11):1334–1345.

[12] Hsieh K, Rimmer JH, Heller T. Obesity and associated factors in adults with intellectual disability. J Intellect Disabil Res. 2014;58(9):851–863.

[13] Rimmer JH, Wang E, Yamaki K, et al. Documenting disparities in obesity and disability (24). Austin (TX): SEDL; 2010. (FOCUS Technical Brief).

[14] Schoufour JD, Mitnitski A, Rockwood K, et al. Development of a frailty index for older people with intellectual disabilities: results from the HA-ID study. Res Dev Disabil. 2013; 34(5):1541–1555.

[15] Oppewal A, Hilgenkamp TIM, van Wijck R, et al. Physical fitness is predictive for a decline in the ability to perform instrumental activities of daily living in older adults with intellectual disabilities: results of the HA-ID study. Res Dev Disabil. 2015;41–42:76–85.

[16] Morin D, Mérineau-Côté J, Ouellette-Kuntz H, et al. A comparison of the prevalence of chronic disease among people with and without intellectual disability. Am J Intellect Dev Disabil. 2012;117(6):455–463.

[17] Oppewal A, Hilgenkamp TIM. Physical fitness is predictive for 5-year survival in older adults with intellectual disabilities. J Appl Res Intellect Disabil. 2019;32(4):958–966.

[18] Pestana MB, Barbieri FA, Vitório R, et al. Effects of physical exercise for adults with intellectual disabilities: a systematic review. J Phys Educ. 2018;29:1–16.

[19] Humphries B, Dugan E, Doyle T. Muscular fitness. In: Medicine ACoS, editor. ACSM's resource manual for guidelines for exercise testing and prescription. Philadelphia (PA): Wolters Kluwer; 2016.

[20] Duchateau J, Hainaut K. Isometric or dynamic training: differential effects on mechanical properties of a human muscle. J Appl Physiol Respir Environ Exerc Physiol. 1984; 56(2):296–301.

[21] Baker D, Wilson G, Carlyon B. Generality versus specificity: a comparison of dynamic and isometric measures of strength and speed-strength. Eur J Appl Physiol Occup Physiol. 1994;68(4):350–355.

[22] Garber CE, Blissmer B, Deschenes MR, et al. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardio-respiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: Guidance for prescribing exercise. Med Sci Sports Exerc. 2011;43(7):1334–1359.

[23] Cowley PM, Ploutz-Snyder LL, Baynard T, et al. The effect of progressive resistance training on leg strength, aerobic capacity and functional tasks of daily living in persons with Down syndrome. Disabil Rehabil. 2011;33(23–24): 2229–2236.

[24] Pitetti KH, Rimmer JH, Fernhall B. Physical fitness and adults with mental retardation. An overview of current research and future directions. Sports Med. 1993;16(1): 23–56.

[25] Terblanche E, Boer P-H. The functional fitness capacity of adults with Down syndrome in South Africa. J Intellect Disabil Res. 2013;57(9):826–836.

[26] Carmeli E, Ariav C, Bar-Yossef T, et al. Movement skills of younger versus older adults with and without Down syndrome. Res Dev Disabil. 2012;33(1):165–171.

[27] Carmeli E, Imam B, Levi R, et al. Hand grip strength is associated with body sway rate among older adults with intellectual disability. J Policy Pract Intellect Disabil. 2013;10(4): 321–325.

[28] Bouzas S, Martínez-Lemos RI, Ayán C. Effects of exercise on the physical fitness level of adults with intellectual disability: a systematic review. Disabil Rehabil. 2019;41(26): 3118–3140.

[29] Shields N, Dodd K. A systematic review on the effects of exercise programmes designed to improve strength for people with Down syndrome. Phys Ther Rev. 2004;9(2): 109–115.

[30] American College of Sports Medicine. ACSM's guidelines for exercise testing and prescription. 9th ed. Philadelphia

[31] Shields N, Taylor NF, Dodd KJ. Effects of a community-based progressive resistance training program on muscle performance and physical function in adults with Down syndrome: a randomized controlled trial. Arch Phys Med Rehabil. 2008;89(7):1215–1220.

[32] Fornieles G, Rosety M, Elosegui S, et al. Salivary testosterone and immunoglobulin A were increased by resistance training in adults with Down syndrome. Braz J Med Biol Res. 2014;47(4):345–348.

[33] Tamse TR, Tillman MD, Stopka CB, et al. Supervised moderate intensity resistance exercise training improves strength in Special Olympic athletes. J Strength Cond Res. 2010; 24(3):695–700.

[34] Hassan NM, Landorf KB, Shields N, et al. Effectiveness of interventions to increase physical activity in individuals with intellectual disabilities: a systematic review of randomised controlled trials. J Intellect Disabil Res. 2019;63(2): 168–191.

[35] Bartlo P, Klein PJ. Physical activity benefits and needs in adults with intellectual disabilities: systematic review of the literature. Am J Intellect Dev Disabil. 2011;116(3):220–232.

[36] Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4(1):1.

[37] Folland JP, Williams AG. The adaptations to strength training: morphological and neurological contributions to increased strength. Sports Med. 2007;37(2):145–168.

[38] Staron RS, Karapondo DL, Kraemer WJ, et al. Skeletal muscle adaptations during early phase of heavy-resistance training in men and women. J Appl Physiol. 1994;76(3): 1247–1255.

[39] Higgins JPT, Thomas J, Chandler J, et al. Cochrane handbook for systematic reviews of interventions. 2nd ed. Hoboken (NJ): John Wiley & Sons Inc.; 2019.

[40] Clarivate Analytics. EndNote. X9.2. 2020.

[41] Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. BMJ. 2005;331(7524):1064–1065.

[42] Cooper H, Hedges LV, Valentine JC. The handbook of research synthesis and meta-analysis. Thousand Oaks (CA): Russell Sage Foundation; 2019.

[43] Smart NA, Waldron M, Ismail H, et al. Validation of a new tool for the assessment of study quality and reporting in exercise training studies: TESTEX. Int J Evid Based Healthcare. 2015;13(1):9–18.

[44] Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20(1):37–46.

[45] Collaboration TC. Review Manager (RevMan). [Computer program]. Copenhagen (Denmark): The Nordic Cochrane Centre, The Cochrane Collaboration; 2011.

[46] Borenstein M, Hedges LV, Higgins JPT, et al. Introduction to meta-analysis. West Sussex (UK): John Wiley & Sons, Ltd; 2009.

[47] Abernethy P, Wilson G, Logan P. Strength and power assessment. Issues, controversies and challenges. Sports Med. 1995;19(6):401–417.

[48] Li RC, Wu Y, Maffulli N, et al. Eccentric and concentric isokinetic knee flexion and extension: a reliability study using the Cybex 6000 dynamometer. Br J Sports Med. 1996;30(2): 156–160.

(PA): Paul D Thompson Wolters Kluwer/Lippincott Williams & Wilkins; 2014.

[49] Higgins JPT, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. The BMJ. 2011;343(2):d5928–d5929.

[50] Rimmer JH, Kelly LE. Effects of a resistance training program on adults with mental retardation. J Hum Kinet. 1991; 8(2):146–153.

[51] Rimmer JH, Heller T, Wang E, et al. Improvements in physical fitness in adults with Down syndrome. Am J Mental Retard. 2004; 109(2):165–174.

[52] Shields N, Bennell KL, Radcliffe J, et al. Is strength training feasible for young people with Prader-Willi syndrome? A phase I randomised controlled trial. Physiotherapy. 2020; 106:136–144.

[53] van Schijndel – Speet M, Evenhuis H, van Wijck R, et al. A structured physical activity and fitness programme for older adults with intellectual disabilities: results of a cluster-randomised clinical trial. J Intellect Disabil Res. 2017; 61(1):16–29.

[54] Suomi R. Self-directed strength training: its effect on leg strength in men with mental retardation. Arch Phys Med Rehabil. 1998;79(3):323–328.

[55] Suomi R, Surburg PR, Lecius P. Effects of hydraulic resistance strength training on isokinetic measures of leg strength in men with mental retardation. Adapt Phys Act Q. 1995;12(4):377–387.

[56] Ko T, Han G, Cho B. Effect of 24 weeks of physical activity therapy on the low leg muscular strength. J Phys Ther Sci. 2012;24(4):351–353.

[57] Calders P, Elmahgoub S, de Mettelinge TR, et al. Effect of combined exercise training on physical and metabolic fitness in adults with intellectual disability: a controlled trial. Clin Rehabil. 2011;25(12):1097–1108.

[58] Oviedo GR, Guerra-Balic M, Baynard T, et al. Effects of aerobic, resistance and balance training in adults with intellectual disabilities. Res Dev Disabil. 2014;35(11):2624–2634.

[59] Jeng S-C, Chang C-W, Liu W-Y, et al. Exercise training on skill-related physical fitness in adolescents with intellectual disability: a systematic review and meta-analysis. Disabil Health J. 2017;10(2):198–206.

[60] Cohen P. To be or not to be: control and balancing of type I and type II errors. Eval Program Plan. 1982;5(3):247–253.

[61] Ptomey LT, Willis EA, Lee J, et al. The feasibility of using pedometers for self-report of steps and accelerometers for measuring physical activity in adults with intellectual and developmental disabilities across an 18-month intervention. J Intellect Disabil Res. 2017;61(8):792–801.

[62] Waters L, Reeves M, Fjeldsoe B, et al. Control group improvements in physical activity intervention trials and possible explanatory factors: a systematic review. J Phys Act Health. 2012;9(6):884–895.

[63] Rodrigues AM, O'Brien N, French DP, et al. The question-behavior effect: genuine effect or spurious phenomenon? A systematic review of randomized controlled trials with meta-analyses. Health Psychol. 2015;34(1):61–78.

[64] Leung W, Siebert EA, Yun J. Measuring physical activity with accelerometers for individuals with intellectual disability: a systematic review. Res Dev Disabil. 2017;67:60–70.

[65] van Schijndel-Speet M, Evenhuis HM, van Wijck R, et al. Facilitators and barriers to physical activity as perceived by older adults with intellectual disability. Intellect Dev Disabil. 2014;52(3):175–186.

[66] American College of Sports Medicine. ACSM's resource manual for Guidelines for exercise testing and prescription. 7th ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2014. English. Retrived from https://www.aaidd.org/docs/default-source/default-document-library/obrusnikova.pdf?sfvrsn=aa13521_0

[67] Obrusnikova I, Cavalier AR, Novak HM, et al. Effects of a community-based familiarization intervention on independent performance of exercise tasks by adults with intellectual disabilities. J Intellect Dev Disabil. in press.

[68] Obrusnikova I, Novak HM, Cavalier AR. The effect of systematic prompting on the acquisition of five muscle-strengthening exercises by adults with mild intellectual disabilities. Adapt Phys Act Q. 2019;36(4):447–471.

[69] Carter KMS, Swank AM. Special considerations for exercise testing and prescription for the individual with an intellectual disability. ACSMs Health Fitness J. 2014;18(2):37–38.

[70] Alexander NB. Postural control in older adults. J Am Geriatr Soc. 1994;42(1):93–108.

[71] Hsieh K, Rimmer JH, Heller T. Prevalence of falls and risk factors in adults with intellectual disability. Am J Intellect Dev Disabil. 2012;117(6):442–454.

[72] Zablotsky B, Black LI, Blumberg SJ. Estimated prevalence of children with diagnosed developmental disabilities in the United States, 2014-2016. NCHS Data Brief, no 291. Hyattsville (MD): National Center for Health Statistics; 2017.

[73] Schalock RL, Borthwick-Duffy SA, Bradley VJ, et al. Intellectual disability: definition, classification, and systems of supports. 11th ed. Washington (DC): American Association on Intellectual and Developmental Disabilities; 2010. English.

[74] Luckasson R, Schalock RL. What's at stake in the lives of people with intellectual disability? Part II: recommendations for naming, defining, diagnosing, classifying, and planning supports. Intellect Dev Disabil. 2013;51(2):94–101.

[75] Grgic J, Lazinica B, Schoenfeld BJ, et al. Test–retest reliability of the one-repetition maximum (1RM) strength assessment: a systematic review. Sports Med Open. 2020;6(1):31.

[76] Schoenfeld BJ, Grgic J, Ogborn D, et al. Strength and hypertrophy adaptations between low- vs. high-load resistance training: a systematic review and meta-analysis. J Strength Cond Res. 2017;31(12):3508–3523.

[77] Brown LE, Weir JP. ASEP procedures recommendation I: accurate assessment of muscular strength and power. J Exerc Physiol Online. 2001;4(3):1–21.

[78] Hilgenkamp TIM, van Wijck R, Evenhuis HM. Feasibility and reliability of physical fitness tests in older adults with intellectual disability: a pilot study. J Intellect Dev Disabil. 2012; 37(2):158–162.

[79] Bocalini DS, Portes L, Ribeiro K, et al. Insight for learning and stability of one repetition maximum test in subjects with or without experience on resistance training. Gazzetta Medica Italiana. 2013;172(11):845–851.

[80] do Nascimento MA, Januário RSB, Gerage AM, et al. Familiarization and reliability of one repetition maximum strength testing in older women. J Strength Cond Res. 2013;27(6):1636–1642.

[81] Griffin L, Cafarelli E. Resistance training: cortical, spinal, and motor unit adaptations. Can J Appl Physiol. 2005;30(3): 328–340.

[82] American College of Sports Medicine. Progression models in resistance training for healthy adults. MedSciSports Exerc. 2009;41(3):687–708.

[83] Obrusnikova I, Cavalier AR, Novak HM, et al. The effect of systematic prompting on the acquisition of two muscle-strengthening exercises by adults with moderate disabilities. J Behav Educ. 2020;29(3):584–605.

[84] Coffey VG, Hawley JA. The molecular bases of training adaptation. Sports Med. 2007;37(9):737–763.

[85] Rintala P, McCubbin JA, Dunn JM. Familiarization process in cardiorespiratory fitness testing for persons with mental retardation. Sports Med Train Rehabil. 1995;6(1):15–27.

[86] Pitetti KH. A reliable isokinetic strength test for arm and leg musculature for mildly mentally retarded adults. Arch Phys Med Rehabil. 1990;71(9):669–672.