JIDR Journal of Intellectual Disability Research
Published on behalf of mencap and in association with IASSID

Journal of Intellectual Disability Research doi: 10.1111/jir.12935

VOLUME PART

Special Issue

The responsiveness of muscle strength tests in adults with intellectual disabilities

S. Weterings,[1,2] A. Oppewal,[1] S. M. A. Bierma-Zeinstra[3] & T. I. M. Hilgenkamp[1,4]

1 *Department of General Practice, Intellectual Disability Medicine, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands*
2 *Abrona, Healthcare Provider for People with an Intellectual Disability, Huis ter Heide, The Netherlands*
3 *Department of General Practice, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands*
4 *Department of Physical Therapy, University of Nevada, Las Vegas, Nevada, USA*

Abstract

Background Muscle strength is both a strong predictor for future negative health outcomes and a prerequisite for physical fitness and daily functioning of adults with ID. Therefore, it is important to be able to monitor the muscle strength of adults with ID over time. The aim of this study is to assess the responsiveness of five field tests that measure muscle strength and endurance (grip strength, hand-held dynamometry of leg extension and arm flexion, 10RM-test of the seated squat and the biceps curl, 30-s chair stand and the 5-times Chair stand) in adults with ID after a 24-week resistance-exercise training (RT) programme.

Method The responsiveness of the five muscle strength and endurance tests was assessed by correlating the change scores of the five tests with the slope of the training progression of specific exercises within the RT-programme, namely, the step up, seated squat, biceps curl and triceps curl.

Results The 10RM-test of the seated squat was significantly correlated with the step up (R = 0.53, P = 0.02) and the seated squat (R = 0.70 P = 0.00). None of change scores on the other tests was significantly correlated with the training progression of the exercises.

Conclusion The 10RM test of the seated squat could potentially be used to evaluate the effects of an RT-programme in adults with ID. Responsiveness of the grip strength, hand held dynamometry, 10RM-test of the biceps curl, 30-s chair stand and the 5-times chair stand could not yet be confirmed.

Keywords adults, intellectual disabilities, muscle strength, progressive resistance training, responsiveness

Correspondence: Mr Stijn Weterings, Department of General Practice, Intellectual Disability Medicine, Erasmus MC University Medical Center Rotterdam, NA1909, P.O. Box 2040, 3000 CA, Rotterdam, The Netherlands (e-mail: s.weterings@erasmusmc.nl).

Background

Adults with intellectual disabilities (ID) generally have lower muscle strength and muscle endurance compared with the general population (Graham and Reid 2000; Hilgenkamp *et al.* 2012; Cuesta-Vargas and Hilgenkamp 2015). Sarcopenia (the age-related loss in muscle mass) is already highly prevalent in adults with ID aged 50–64 years old (Bastiaanse *et al.* 2012; Carmeli *et al.* 2012). Lower muscle strength has been found to be predictive for a decline in the performance of both basic and instrumental activities of daily living (ADL), for a decline in mobility and for a higher mortality risk in adults with ID (Oppewal *et al.* 2014; Oppewal and Hilgenkamp 2019b). With muscle strength being both a strong predictor for future negative health outcomes, and a prerequisite for physical fitness and daily functioning of adults with ID, it is important to be able to monitor the muscle strength of adults with ID over time.

Monitoring muscle strength and muscular endurance requires exercise tests with good measurement properties, such as feasibility, reliability, validity and responsiveness. Tests normally used in the general population cannot simply be used in adults with ID, because adults with ID often have motivational, behavioural and/or physical problems, as well as cognitive limitations that hamper the execution of some tests (Bossink *et al.* 2017; Riebe *et al.* 2018). It requires the expertise of trainers to motivate the participants to exercise/work out/perform at their best, but even then it can sometimes be difficult (Weterings *et al.* 2020a). Previous studies with adults with ID have used different tests to measure muscle strength and muscle endurance, such as a one-repetition-maximum (estimation) test (1RM-test) (Machek *et al.* 2008; Shields and Taylor 2010; Calders *et al.* 2011; Mendonca *et al.* 2011; Dijkhuizen *et al.* 2018), the maximal voluntary contraction test measured with a hand-held dynamotor (HHD) (Lin and Wuang 2012), the grip strength (GS) test measured with a hand dynamometer (Hilgenkamp *et al.* 2012; van Schijndel-Speet *et al.* 2016), the 30-s chair stand test (30sCS) (Podgorski *et al.* 2004; Hilgenkamp *et al.* 2012; Dijkhuizen *et al.* 2018) and the 5-times chair stand test (5tCS) (used by the Healthy Athletes programme of the Special Olympics) (Oppewal and Hilgenkamp 2019a) to measure muscle endurance. All these tests have been found feasible and reliable in measuring muscle strength or endurance in adults with ID (Dunn 1978; Surburg *et al.* 1992; Horvat *et al.* 1993; Dijkhuizen *et al.* 2018; Oppewal and Hilgenkamp 2019a) and the GS and 30sCS are also predictive for a decline in mobility (Oppewal and Hilgenkamp 2019a). However, the responsiveness of all these muscle strength and endurance tests is unknown.

Responsiveness is defined as 'the ability of an instrument or test to detect change over time in the construct to be measured' (Mokkink *et al.* 2010). Assessing whether an individual's status has changed over time is often the most important objective of measurements in clinical practice and research (Vet *et al.* 2015). If a test is not responsive, it cannot determine whether the muscle strength of a group or an individual changed over time.

Therefore, the aim of this study is to assess the responsiveness of five field tests that measure muscle strength and endurance (GS, HHD, 10RM-tests, 30sCS and 5tCS) in adults with ID. We will assess the responsiveness of these tests over a 24-week resistance-exercise training programme (RT-programme). It is expected that there will be a significant positive correlation between the change scores of the five field tests and the progression (change scores) on the exercises (step up, seated squat, biceps curl, triceps curl) of this 24-week resistance-exercise training programme.

Method

Study design

This study was part of a multicentre feasibility study of vigorous resistance-exercise training for adults with ID with cardiovascular disease (CVD) risk factors (Weterings *et al.* 2020a), which was conducted by the 'Healthy Ageing and Intellectual Disabilities' (HA-ID) consortium. This consortium consists of three care provider organisations for people with ID in the Netherlands, Abrona (Huis ter Heide), Ipse de Bruggen (Zoetermeer) and Amarant (Tilburg), and the Chair for Intellectual Disability Medicine of the Erasmus MC, University Medical Center Rotterdam (Hilgenkamp *et al.* 2011).

Participants

The participants lived and/or worked in a residential or community-based setting of the participating care provider organisations. They were invited to participate in the RT-programme by their nurse practitioner if they were diagnosed with at least one CVD risk factor (type 2 diabetes-mellitus, hypertension, hypercholesterolemia and/or overweight/obesity), were 18 years or older, had a mild (intelligence quotient (IQ) = 50–69) or moderate (IQ = 35–49) ID, and when a training facility was present nearby. Participants were excluded when physical problems inhibited exercising or when their physician had not provided medical clearance (Weterings *et al.* 2020a). This study was performed in accordance with the declaration of Helsinki (WMA 2013) and all participants or their legal representatives provided written informed consent. The medical ethics committee of the Erasmus MC, University Medical Center Rotterdam, the Netherlands, approved this study (MEC-2016-574). All participants who finished the RT-programme were included in the analysis for this study.

Overview study procedures

The participants completed a 24-week RT-programme, with two training sessions a week (Weterings *et al.* 2020a). Each session consisted of seven exercises (step up, seating squat, abdominal curl, bridge pose, biceps curl and triceps curl), which were found feasible for adults with ID (Weterings *et al.* 2018). A physiotherapist or physical activity instructor with experience with working with adults with ID supervised all sessions. The RT-programme consisted of five training intensity phases. The training intensity was defined by the percentage of 1RM (%1RM). The 1RM is the maximum amount of weight that a person can possibly lift for one repetition over the whole range of motion (Riebe *et al.* 2018). The RT-programme started with a familiarisation phase at 55% of 1RM (2 series of 20 repetitions), then a training phase at 60% of 1RM (2 series of 18 repetitions), then at 70% of 1RM (3 series of 12 repetitions), 75% of 1RM (3 series of 10 repetitions) and finally a training phase at 80% of 1RM (3 series of 8 repetitions) (Weterings *et al.* 2020a). For each training session, the trainers logged the weight and the number of repetitions for all series of all exercises. For the step up the height of the step that was used was logged as well.

At baseline, the participants performed the GS, the HHD of the arms and legs, the 30sCS and the 5tCS. After the familiarisation phase, the 10RM-test was performed for the seated squat and the biceps curl. The 10RM-tests were performed after the familiarisation phase to assure safety and a good execution of the test. The duration of the familiarisation phase differed for all participants and ended when the RT-exercises were performed with good posture and technique (see description below) for eight consecutive sessions (Weterings *et al.* 2020a). At the end of the RT-programme all the strength and muscular endurance tests were repeated within a 2- to 5-day interval after the last training session. The test administrator was a physiotherapist with 15 years of experience in working with adults with ID performed all measurements (S.W.). During all measurements the participants were verbally motivated as much as possible.

Measurements

Participant characteristics

Age, sex, level of ID, CVD risk factors (type 2 diabetes, hypertension, dyslipidaemia and/or overweight/obesity), and diagnosis of Down syndrome and cerebral palsy were collected from medical records. Body mass index (BMI) was calculated by dividing weight (measured with Seca Robusta type 813, in kilogram) by squared height (measured with Seca 216 height rod, in metre). Waist circumference was measured with a non-stretchable measurement tape over the unclothed abdomen at the narrowest point between the costal margin and iliac crest in a standing position with the arms along the body (in centimetres).

Muscle strength tests

Grip strength test. The GS was measured with a hand dynamometer (Jamar hand dynamometer, Sammons Preston Rolyan, Bolingbrook, IL, USA). The participant was sitting in a chair with the elbow in a 90-degree angle and the hand palm in a vertical position. The test was performed three times for both hands with 1-min rest between the attempts of the same hand. The maximum score of the six attempts was the test score (Oppewal and Hilgenkamp 2019a).

Hand-held dynamometry test. The maximal voluntary contraction of both the arms and legs was measured with a handheld dynamometer (HHD-test) (Microfet 2, Hoggan Health Industries) for both the flexion and extension movement, using the break-method

(Bohannon 1988; Burns and Spanier 2005). The HHD measurements of the elbow flexion and extension were performed sitting behind a table with the elbow resting on a table at a 90-degree angle. Knee flexion and extension measurements were performed in a prone position with both hands resting under the head. The knee was placed vertical in a 90-degree angle. The HHD placement was at the most distal point of the lower arm and leg (van der Ploeg *et al.* 1984). The HHD-test was found feasible and reliable in adults with ID (Weterings *et al.* 2020b). In the general population, feasibility, reliability and the sensitivity to change were also good (van der Ploeg *et al.* 1991; Bohannon 1997).

10RM-test. In this study, the participants performed the 10RM-test of the seated squat (10RM-Seated squat) and the biceps curl (10RM-Biceps curl). We used the 10RM-test instead of the 1RM-test, because of the higher risk for injury with the 1RM-test in frail groups, like elderly and chronically ill people (Garber *et al.* 2011). The 10RM-test has been found feasible and reliable for people with high risk of CVD and health conditions in the general population (Williams *et al.* 2007). Additionally, the 10RM-test seemed more suited for adults with ID, because they are often not used to perform vigorous exercises (Bossink *et al.* 2017). The weights for the 10RM-tests were selected in consultation with the trainer, choosing the weight for which the participant was expected to be able to perform 8–12 repetitions. With that weight, the participant was asked to perform the exercise until exhaustion, with a maximum of 30 repetitions. When a participant reached 30 repetitions, he or she was asked to stop and perform the 10RM-test again with a higher weight after a resting period of at least 5 min. The number of performed repetitions was then used to estimate at which percentage of 1RM the test was performed (see Table 1 for the percentages). The weight and the percentage of 1RM were used to calculate the score for the 10RM-test. For example, a participant performed 14 repetitions with 12 kg for the biceps curl; 14 repetitions can be compared with 70% of 1RM so the score for the 10RM-Biceps curl = 12 kg*(100/70) = 17.1 kg.

Table 1 The number of repetitions and presented as the percentage of 1RM (Jongert *et al.* 2004; Garber *et al.* 2011)

No. of repetitions	Percentage of 1RM
6–8	80
9–11	75
12–14	70
15–16	65
17–19	60
20–24	55
25–30	50

1RM, one repetition maximum; no., number.

Muscle endurance tests

Chair stand tests. For the 30-s chair stand test (30sCS), participants were instructed to stand up and sit down again as fast as possible in 30 s, without using their hands. The number of repetitions was the test score. For the 5-times chair stand test (5tCS), participants were instructed to stand up and sit down again as fast as possible for five times, without using their hands. The time to complete five stances was the test result. The test administrator recorded time at one-hundredth of a second with a stopwatch. In both tests, the participants started sitting on a chair with the knees in a 90-degree angle and the feet on the floor (Oppewal and Hilgenkamp 2019a).

Training progression

The responsiveness of the five muscle strength and endurance tests was assessed by correlating the change scores of the five tests with the slope of the training progression of the step up, seated squat, biceps curl and/or triceps curl. We selected these specific exercises from the total of seven exercise performed within the RT-programme, because the tests measure the muscle strength and endurance of the muscle groups targeted with these specific exercises.

To determine the progression in the training sessions, we calculated the average slope. The 1RM-score of each exercise of all training session was estimated. We used the number of performed repetitions to determine the training intensity of each exercise and training session (Table 2), which was then used to calculate the 1RM-score (see below for the specific calculations for each exercise). All the 1RM-scores were used as data points to create a slope

Table 2 Participants characteristics

Number of participants, *n*	19
Male, *n* (%)	10 (45.8%)
Female, *n* (%)	9 (54.2%)
Level of ID	
Mild, *n* (%)	9 (45.8%)
Moderate, *n* (%)	10 (54.2%)
Diagnoses	
Down syndrome, *n* (%)	3 (12.5%)
Cerebral palsy (GMFCS I), *n* (%)	2 (9.5%)
Age (in years), mean ± SD [range]	42 ± 18 [23–75]
CVD risk factors	
Type 2 diabetes mellitus, *n* (%)	5 (29%)
Hypertension, *n* (%)	6 (29%)
Dyslipidaemia, *n* (%)	4 (20%)
Overweight/obese, *n* (%)	16 (92%)
BMI mean ± SD [range]	33.9 ± 6.9 [17.4–44.2]
Waist circumference (in cm), mean ± SD [range]	114 ± 14 [82–139]

ID, intellectual disability; GMFCS, Gross Motor Function Classification Score; SD, standard deviation; CVD, cardiovascular disease; BMI, body mass index; cm, centimetre.

of the training progression throughout the programme of each participant for each exercise.

Step up

The 1RM-score of the step-up was calculated by multiplying the total training weight (bodyweight plus added training weight in kilogram) with the height of the step (in metre). This result was then multiplied with 100 divided by the training intensity percentage to obtain the 1RM-score of the step-up in kilogram*metre: (the 1RM score = ((body weight + training weight)*height*100)/%1RM) (Zatsiorsky *et al.* 2021).

Seated squat

For the seated squat, 80% of the bodyweight is used during training (de Leva 1996). Therefore, the 1RM-score was calculated by multiplying the total weight (0.8 times bodyweight plus added training weight) with a 100 divided by the training intensity percentage to obtain the 1RM-score in kilogram. Some participants used a leg press. For them, the estimated-1RM was calculated by multiplying added training weight with a 100 divided by the training intensity percentage to obtain the 1RM-score in kilogram.

Biceps curl and triceps curl

For both the biceps and triceps curl, the estimated-1RM was calculated by multiplying training weight (the total weight held in the left and right arm) with a 100 divided by the training intensity percentage to obtain the 1RM-score in kilogram.

Statistical analyses

The participant's characteristics were analysed with descriptive statistics for all participants who finished the RT-programme. The results of the muscle strength and endurance tests were analysed with descriptive statistics and a *t*-test to test for differences between before and after the RT-programme. The results of the 1RM-scores of each exercise of the first and last training session were analysed with descriptive statistics, a paired samples *t*-test to test for differences and the effect sizes (ES) were calculated with Cohen's *d*. ES were considered low (<0.2), moderate (>0.2 to <0.8) or large (>0.8) (Cohen 1988).

To assess the responsiveness of the tests, a linear mixed model (LMM) was used to compare the slope of the 1RM-score on the exercises (training progression) of each participant with the change scores of the muscle strength and endurance tests for each participant. The assumptions of normality were checked and were not perfectly met for some data; however, this was considered to not influence our model fit. The responsiveness for each test was calculated by averaging the individual correlation scores between the tests and the training progression of the exercises. We calculated the individual correlations because we anticipated that if the results were pooled, some correlations could disappear because of the large heterogeneity of the participants and their training results.

For the lower extremities, the change scores of the 30sCS, the 5tCS and the HHD-test of the legs were compared with the slope of the training progression on the step up and seated squat, from the start until the end of the RT-programme. Furthermore, the change scores of the 10RM-Seated squat were compared with the slope of the training progression of the step up and seated squat starting after the familiarisation phase until the end of the RT-programme.

For the upper extremities, the change scores of the GS and the HHD-test of the arms were compared with the slope of the training progression of the biceps curl and the triceps curl at the start and end of the RT-programme. Furthermore, the change scores of the 10RM-Biceps curl were compared with the training progression of the biceps curl starting after the familiarisation phase until end of the RT-programme.

During the training, some participants changed from a seated squat to a leg press. When this happened, we saw that the estimated-1RM of the seated squat showed a sudden drop. This did not resemble the actual training experiences of the participants and trainers. It is most likely caused by the formula used to calculate the estimated-1RM, by either overestimating the amount of weight lifted by the seated squat or by underestimating the force needed during the leg press performance. Therefore, we corrected for this sudden drop in the LMM by equalising the first 1RM-score of the leg press with the last 1RM-score of the seated squat and progressing from that point on. All LMM analyses were performed in R Studio (R Studio, Boston), and the descriptive statistics and *t*-tests were analysed in SPSS 25 (IBM corporation).

Results

Participant's characteristics

Nineteen participants (10 men and 9 women) out of 24 (12 men and 13 women) finished the RT-programme and were included in the analyses. Nine participants had a mild ID, and 10 had a moderate ID. Five participants had type 2 diabetes mellitus, six

Table 3 Results of the muscle strength and endurance tests at baseline and post-intervention

Exercise test	Baseline measurement, mean ± SD	Post-intervention measurement, mean ± SD	Change score	Effect size Cohen's *d* [confidence interval]	*P*-value
GS (in kg)	35.0 ± 9.5	35.0 ± 11.0	0.0 ± 3.5	0.00 [−0.64 to 0.64]	1.00
HHD elbow flexion left (in N)	215 ± 69	233 ± 80	18 ± 33	0.24 [−0.4 to 0.87]	0.04*
HHD elbow flexion right (in N)	207 ± 74	227 ± 80	19 ± 25	0.26 [−0.38 to 0.89]	0.00**
HHD elbow extension left (in N)	150 ± 47	154 ± 56	4 ± 25	0.08 [−0.56 to 0.71]	0.51
HHD elbow extension right (in N)	147 ± 40	156 ± 54	10 ± 25	0.21[−0.45 to 0.82]	0.10
HHD knee flexion left (in N)	158 ± 58	150 ± 55	−7 ± 29	−0.14 [−0.78 to 0.50]	0.30
HHD knee flexion right (in N)	167 ± 58	156 ± 54	−11 ± 27	−0.20 [−0.83 to 0.45]	0.08
HHD knee extension left (in N)	264 ± 110	267 ± 109	2 ± 36	0.03 [−0.61 to 0.66]	0.80
HHD knee extension right (in N)	270 ± 107	268 ± 104	3 ± 46	−0.02 [−0.65 to 0.62]	0.78
10RM-Seated squat (1RM in kg) after familiarisation	166.5 ± 60.4	203.2 ± 89.8	36.7 ± 73.8	0.48 [−0.18 to 1.11]	0.04*
10RM-Biceps curl (1RM in kg)+ after familiarisation	18.3 ± 6.7	23.4 ± 9.5	5.3 ± 5.9	0.62 [−0.04 to 1.26]	0.00**
30sCS (no.)	13.7 ± 4.8	14.7 ± 6.2	−1.1 ± 3.5	0.18 [−0.46 to 0.81]	0.17
5tCS (s)	11.53 ± 6.12	11.17 ± 4.88	−0.36 ± 5.89	−0.07 [−0.70 to 0.50]	0.79

Paired *t*-test.

*$P < 0.05$.

**$P < 0.01$.

+Based on 18 participants; one participant could not perform the 10RM-Biceps curl at the end of the programme.

SD, standard deviation; HHD, maximal voluntary contraction measured with a hand held dynamometer; N, Newton; RM, repetition maximum; GS, grip strength; 30sCS, 30-s chair stand; 5tCS, five-times chair stand; no., number of repetitions.

had hypertension, four had dyslipidaemia and 16 were overweight or obese (Table 2).

The muscle strength tests

The results of the muscle strength tests are shown in Table 3. The duration of the familiarisation phase differed per participant and had an average of 22.1 (±9.5) sessions (ranged between session 10 and session 46).

The training progression of the exercises

The average 1RM-scores of the first training session and the last training session of the participants of each exercise are shown in Table 4.

The responsiveness

The correlations between the training progression of the exercises and the change score on the tests are shown in Table 5. The 10RM-Seated squat was significantly correlated with the step up (R = 0.53, P = 0.02) and the seated squat (R = 0.70 P = 0.00). Furthermore, the HHD knee extension of the right leg was significant negatively correlated with the step up (R = −0.52, P = 0.02) but not significant with the seated squat (P = 0.80). None of change scores on the other tests were significantly correlated with the training progression on the other exercises. In Figure 1, an example of the training progression of the seated squat for each participant is shown. For each participant, the 1RM-score of each training session is plotted; the line represents the average slope of training progression of the whole RT-programme. This slope is correlated with the change score of the tests for each participant.

Discussion

For this study, we assessed the responsiveness of the GS, HHD-test, 10RM-tests, 30sCS and 5tCS by correlating the changes in performance on these tests with the slope of the progression of the exercises of a 24-week RT-programme in adults with ID. In this study, only the 10RM-Seated squat seems to be a responsive test to measure the progress in muscle strength over the RT-programme.

All four exercises within the RT-programme showed a significant progression, with 150% and 230% for the biceps and triceps curl respectively (effect sizes of 1.83 and 2.13). The progression of the step up (35%) and seated squat (13%) were smaller (effect sizes of 0.87 and 0.5). These effect sizes are larger than the effect size of 0.26 Shields et al. (2008) found for the leg press after a 10-week RT-programme for adults with DS (Shields *et al.* 2008). A study by Calders *et al.* (2011) showed a similar progression of 33% for the lower body strength in adults with ID after a 20-week RT-programme (Calders *et al.* 2011).

The 10RM-Seated squat (change score of 36.7 kg) and the 10RM-Biceps curl (change score of 5.3 kg.)

Table 4 The 1RM-scores of the exercises of the RT-programme at the first and last training session

	Group average 1RM-score of first training session	Group average 1RM-score of the last training session	Progression (in %)	Effect size Cohen's *d* [confidence interval]	*P*-value
Step up (in kgm)	35.0	51.0	35%	0.87 [−0.075 to 1.806]	0.001**
Seated squat (in kg)	172.3	195.0	13%	0.5 [−0.413 to 1.413]	0.03*
Biceps curl (in kg)	8.0	20.0	150%	1.83 [0.754 to 2.895]	0.000**
Triceps curl (in kg)	5.3	17.5	230%	2.18 [1.046 to 3.318]	0.000**

Paired *t*-test.
*$P < 0.05$.
**$P < 0.01$.
RM, repetition maximum; %, percentage; kgm, kilogram*metre; kg, kilogram.

Table 5 Correlations between the change scores of the muscle strength tests and the average progression of the 1RM-scores of the exercises

Test	Exercise	Correlation mean [confidence interval]	*P*-value
GS	Biceps curl	−0.35 [−0.69 to 0.13]	0.14
	Triceps curl	0.14 [−0.33 to 0.56]	0.56
HHD elbow flexion left	Biceps curl	−0.19 [−0.61 to 0.30]	0.44
HHD elbow flexion right	Biceps curl	−0.11 [−0.54 to 0.36]	0.65
HHD elbow extension left	Triceps curl	−0.05 [−0.50 to 0.41]	0.82
HHD elbow extension right	Triceps curl	0.00 [−0.46 to 0.45]	0.99
HHD knee extension left	Step up	−0.37 [−0.71 to 0.10]	0.12
	Seated squat	−0.06 [−0.50 to 0.41]	0.82
HHD knee extension right	Step up	−0.52 [−0.79 to −0.09]	0.02*
	Seated squat	−0.06 [−0.50 to 0.40]	0.80
10RM-Seated squat+	Step up	0.53 [0.10 to 0.79]	0.02*
	Seated squat	0.71 [0.31 to 0.88]	0.00**
10RM-Biceps curl+	Biceps curl	−0.45 [−0.76 to 0.03]	0.06
30sCS	Step up	−0.11 [−0.54 to 0.36]	0.65
	Seated squat	0.17 [−0.30 to 0.58]	0.47
5tCS	Step up	0.03 [−0.43 to 0.48]	0.90
	Seated squat	0.05 [−0.41 to 0.49]	0.84

Paired *t*-test.
*$P < 0.05$.
**$P < 0.01$,
+After familiarisation.
30sCS, 30-s chair stand; 5tCS, five-times chair stand; HHD, maximum voluntary contraction measured with hand held dynamometer.

showed a significant improvement. Another study in young adults with ID by Machek et al. (2008) found an increase in the predicted-1RM of seated dip of 53.74 kg and an increase of 25.6 kg for the biceps curl after a 12-week RT-programme (Machek *et al.* 2008). Furthermore, we found a significant improvement in the HHD elbow flexion test (both left and right) with change scores of 18 and 19 N. The HHD knee extension showed no significant improvement. In contrast, a study by Lin *et al.* (2012) found a significant progression in the HHD knee extension test with a change score of 3.42 Pounds after a 6-week RT-programme in adolescents with Down syndrome (Lin and Wuang 2012). The scores of the other HHD tests, the GS, 30sCS and 5tCS did not change in our study. Calders et al. (2011) did find significant improvements for the 30sCS and GS after a 20-week RT-programme in adults with a mild ID (Calders *et al.* 2011).

We hypothesised that the heterogeneity of the participants in our study might impair finding significant results. For example, using the average of the correlations might eliminate differences between responders and non-responders of the RT-programme. Therefore, we used the individual correlations of each participant instead and calculated the average of those individual correlations. Nevertheless, for most tests we found non-significant low correlations. There are very few articles on the responsiveness of muscle strength tests in general, so it is difficult to compare the results of this study with other literature than the previous-mentioned intervention studies. These studies mostly mention the progression of their RT-programme or mention the results on the tests, but no study mentions the correlation between the results of the RT-programme and the results of the tests. We therefore have to hypothesise what could be the reason why the responsiveness of the tests in this study showed mostly low non-significant correlations. There are some potential reasons for this lack of responsiveness.

First, it could be that most of these tests are not so responsive to measure changes in muscle strength and therefore less suited to be used to evaluate the progression of muscle strength within an RT-programme. Responsiveness studies focus on the agreement between change scores, in our case two measurements of the tests and the slope of the

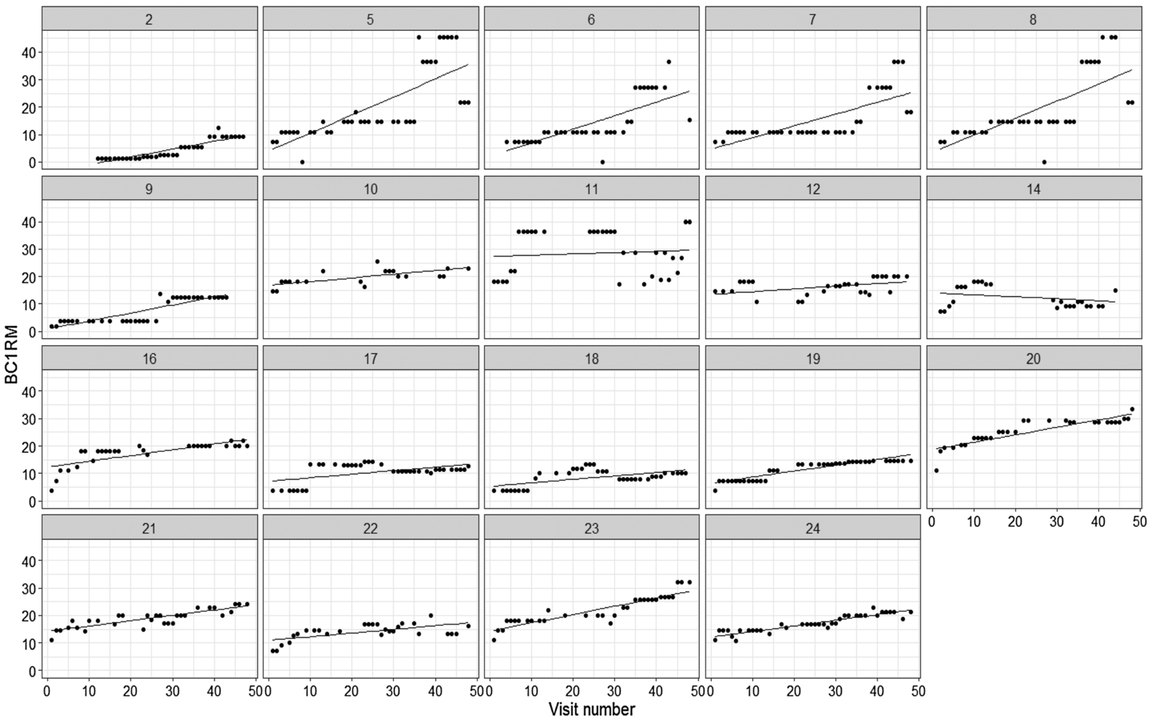

Figure 1. The estimated 1RM of each training session of each participant of the biceps curl with the average slope of the training progression. This example is representative for the other exercises.

exercises, with its own measurement error (often indicated by the MDC of a test) for each measurement or calculation. The MDC of a single score is large for the GS (6.5–8 kg. (Blankevoort *et al.* 2013, Kim *et al.* 2014), the HHD-test (10–17 N.) (Buckinx *et al.* 2017) and the chair stand (two repetitions) (Hesseberg *et al.* 2015) in the general (older) population. These large MDC's lower the potential correlations for their responsiveness, as explained above. It could, therefore, be difficult for these tests to show progression after the RT-programme, despite the significant increase in 1RM-scores in the training programme. In this study, only the responsiveness of the 10RM-Seated squat was significantly correlated with the step up and seated squat. The correlations of the 10RM-Seated squat were 0.5 for the step up and 0.7 for the seated squat. There are, to date, no guidelines on what is an acceptable correlation for responsiveness (Vet *et al.* 2015). Normally, 0.7 is the minimum correlation to be acceptable, but in responsiveness studies lower scores are often found (Vet *et al.* 2015).

Second, it could also be argued that the GS, HHD, 30sCS and 5tCS tests are not suited to measure the increased muscle strength, because these tests measure different aspects of muscle performance. The GS measures strength in the hands and that is not trained specifically. The 30sCS and 5tCS measure muscle endurance. These two tests are highly dependent of the speed with which the test is performed. It could be speculated that though the leg muscles were potentially getting stronger, it was still difficult for adults with ID to speed up the sitting and standing, because speed requires another type of muscle control, coordination and cognitive attention, which can be difficult for adults with ID (Riebe *et al.* 2018). The HHD measures isometric muscle strength of the arms and legs in a 90-degree angle, but the exercises of the RT-programme are performed over a full range of motion of the muscles. Training and measuring muscle strength is dependent on the angle in which it is performed and trained (Riebe *et al.* 2018). Only the 10RM-tests are performed exactly like some of the exercises of the RT-programme. They are the only ones that partly show a correlation, even though they were not performed at the start of the programme.

Limitations

This study was performed within a feasibility study regarding the feasibility of a progressive RT-programme in adults with ID with CVD risk factors. This study sample is not a representative sample of the whole, diverse population of adults with ID and therefore more research into the responsiveness of muscle strength tests is necessary in adults with ID. With 19 participants, this study included just a small heterogeneous sample of adults with ID with at least one CVD risk factor. The heterogeneous sample is preferred for feasibility testing, as it reflects the differences of the adults with ID in daily life, but a heterogeneous sample lowers the internal validity of a study. So the results of this study should be interpreted with caution.

Furthermore, the modelling of the training progression by the 1RM-scores of the exercises in RT-programme could have impeded the actual progression of the participants' muscle strength. The training intensity differed during the RT-programme for most participants and exercises and even exercise execution differed between participants and even within the RT-programme of individual participants. All these different factors made it more complex to model the training results into a standardised 1RM-score. Furthermore, this modelling was based on assumptions when calculating the 1RM-scores, and this could have impeded with the true training progression of the participants in the RT-programme.

Important factors that could influence both the testing results and the progression during the RT-programme were the motivational, behavioural and/or physical problems, and cognitive limitations of the participants (Bossink *et al.* 2017; Riebe *et al.* 2018). It requires the expertise of the trainers to motivate the participants to train at their best, but even then it is hard to interpret if they actually performed the exercises and tests as best as they could. For example, sometimes a participant stopped training after one or two lifts after a training weight was increased (always in small steps), stating this was way too heavy to lift, where the participant easily performed the required 10–20 repetitions the series before. This could also be a problem in the general population but even more so for adults with ID.

Recommendations

There is a need for more uniform measurements with good measurement properties (Robertson *et al.* 2017). This is the first study into the responsiveness of muscle strength tests in adults with ID. More studies are necessary to find the appropriate muscle strength tests to monitor changes in muscle strength of adults with ID. The ID-fitscan is a first attempt to obtain uniformity in fitness tests for adults with ID, which have been found reliable and valid (Oppewal and Hilgenkamp 2019a). However, the responsiveness of the muscle strength and endurance tests used in the ID-fitscan (GS, 30sCS and 5tCS) is still questionable, which could make them unfit to evaluate RT-programmes in adults with ID. Only the responsiveness of the 10RM-Seated squat showed a significant correlation of 0.53 and 0.70 and with the step up and seated squat. Future research should also study potential individual factors influencing the responsiveness of the tests, as there are large differences between participants in the test results and the change scores of the RT-programme as shown in Figure 1. Furthermore, more research is needed into the floor and ceiling effects and smallest detectable change, which have still never been investigated, and are necessary to interpret the results of muscle strength testing.

Conclusion

The 10RM-Seated squat could potentially be used to evaluate the effects of a RT-programme in adults with ID with CVD risk factors. Furthermore, it is still questionable that the GS, HHD-test, 10RM-Biceps curl, 30sCS and 5tCS could be used to evaluate the effects of a RT-programme in adults with ID. Interestingly, the 1RM-scores of all four exercises, both the 10RM tests and the HHD-test of the elbow flexion were all significantly improved. This stresses the need for more research into the interpretation of the results of RT-programmes and the way muscle strength can be measured in adults with ID.

Acknowledgments

The authors would like to thank the participants, the physical activity instructors and physiotherapists of the HA-ID care providers.

Source of funding

This study was carried out with the financial support of the participating care organisations in the Netherlands.

Conflict of interest

The authors declare no conflict of interest.

Ethics committee

The medical ethics committee of the Erasmus Medical Centre at Rotterdam, the Netherlands (approval number: MEC-2016-574).

Data availability statement

Generated data or analysed data from the current study are available from the corresponding author on reasonable request.

References

Bastiaanse L. P., Hilgenkamp T. I. M., Echteld M. A. & Evenhuis H. M. (2012) Prevalence and associated factors of sarcopenia in older adults with intellectual disabilities. *Research in Developmental Disabilities* **33**, 2004–12.

Blankevoort C. G., van Heuvelen M. J. & Scherder E. J. (2013) Reliability of six physical performance tests in older people with dementia. *Physical Therapy* **93**, 69–78.

Bohannon R. W. (1988) Make tests and break tests of elbow flexor muscle strength. *Physical Therapy* **68**, 193–4.

Bohannon R. W. (1997) Reference values for extremity muscle strength obtained by hand-held dynamometry from adults aged 20 to 79 years. *Archives of Physical Medicine and Rehabilitation* **78**, 26–32.

Bossink L. W. M., van der Putten A. A. & Vlaskamp C. (2017) Understanding low levels of physical activity in people with intellectual disabilities: a systematic review to identify barriers and facilitators. *Research in Developmental Disabilities* **68**, 95–110.

Buckinx F., Croisier J. L., Reginster J. Y., Dardenne N., Beaudart C., Slomian J. *et al.* (2017) Reliability of muscle strength measures obtained with a hand-held dynamometer in an elderly population. *Clinical Physiology and Functional Imaging* **37**, 332–40.

Burns S. P. & Spanier D. E. (2005) Break-technique handheld dynamometry: relation between angular velocity and strength measurements. *Archives of Physical Medicine Rehabilitation* **86**, 1420–6.

Calders P., Elmahgoub S., de Mettelinge T. R., Vandenbroeck C., Dewandele I., Rombaut L. *et al.* (2011) Effect of combined exercise training on physical and metabolic fitness in adults with intellectual disability: a controlled trial. *Clinical Rehabilitation* **25**, 1097–108.

Carmeli E., Imam B. & Merrick J. (2012) The relationship of pre-sarcopenia (low muscle mass) and sarcopenia (loss of muscle strength) with functional decline in individuals with intellectual disability (ID). *Archives of Gerontology and Geriatrics* **55**, 181–5.

Cohen J. (1988) *Statistical power analysis for the behavioral sciences*. Erlbaum Associates, Hillsdale, N.J., L.

Cuesta-Vargas A. & Hilgenkamp T. (2015) Reference values of grip strength measured with a Jamar dynamometer in 1526 adults with intellectual disabilities and compared to adults without intellectual disability. *PLoS ONE* **10**, e0129585.

de Leva P. (1996) Adjustments to Zatsiorsky-Seluyanov's segment inertia parameters. *Journal of Biomechanics* **29**, 1223–30.

Dijkhuizen A., Douma R. K., Krijnen W. P., van der Schans C. P. & Waninge A. (2018) Measuring quadriceps strength in adults with severe or moderate intellectual and visual disabilities: feasibility and reliability. *Journal of Applied Research in Intellectual Disabilities* **31**, 1083–90.

Dunn J. M. (1978) Reliability of selected psychomotor measures with mentally retarded adult males. *Perceptual and Motor Skills* **46**, 295–301.

Garber C. E., Blissmer B., Deschenes M. R., Franklin B. A., Lamonte M. J., Lee I. M. *et al.* (2011) American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Medicine and Science in Sports and Exercise* **43**, 1334–59.

Graham A. & Reid G. (2000) Physical fitness of adults with an intellectual disability: a 13-year follow-up study. *Research Quarterly for Exercise and Sport* **71**, 152–61.

Hesseberg K., Bentzen H. & Bergland A. (2015) Reliability of the senior fitness test in community-dwelling older people with cognitive impairment. *Physiotherapy Research International* **20**, 37–44.

Hilgenkamp T. I. M., Bastiaanse L. P., Hermans H., Penning C., van Wijck R. & Evenhuis H. M. (2011) Study healthy ageing and intellectual disabilities: recruitment and design. *Research in Developmental Disabilities* **32**, 1097–106.

Hilgenkamp T. I. M., van Wijck R. & Evenhuis H. M. (2012) Low physical fitness levels in older adults with ID: results of the HA-ID study. *Research in Developmental Disabilities* **33**, 1048–58.

Horvat M., Croce R. & Roswal G. (1993) Magnitude and reliability of measurements of muscle strength across trials for individuals with mental retardation. *Perceptual and Motor Skills* **77**, 643–9.

Jongert T., Koers H. & Oudhof J. (2004) Fitnessinstructie hartfalen (in Dutch), Nederlandse Hartstichting.

Kim J. K., Park M. G. & Shin S. J. (2014) What is the minimum clinically important difference in grip strength? *Clinical Orthopaedics and Related Research* **472**, 2536–41.

Lin H. C. & Wuang Y. P. (2012) Strength and agility training in adolescents with Down syndrome: a randomized controlled trial. *Research in Developmental Disabilities* **33**, 2236–44.

Machek M. A., Stopka C. B., Tillman M. D., Sneed S. M. & Naugle K. E. (2008) The effects of a supervised resistance-training program on Special Olympics athletes. *Journal of Sport Rehabilitation* **17**, 372–9.

Mendonca G. V., Pereira F. D. & Fernhall B. (2011) Effects of combined aerobic and resistance exercise training in adults with and without Down syndrome. *Archives of Physical Medicine and Rehabilitation* **92**, 37–45.

Mokkink L. B., Terwee C. B., Patrick D. L., Alonso J., Stratford P. W., Knol D. L. *et al.* (2010) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research* **19**, 539–49.

Oppewal A., Hilgenkamp T. I., van Wijck R., Schoufour J. D. & Evenhuis H. M. (2014) Physical fitness is predictive for a decline in daily functioning in older adults with intellectual disabilities: results of the HA-ID study. *Research in Developmental Disabilities* **35**, 2299–315.

Oppewal A. & Hilgenkamp T. I. M. (2019a) Adding meaning to physical fitness test results in individuals with intellectual disabilities. *Disability and Rehabilitation*, 1–8.

Oppewal A. & Hilgenkamp T. I. M. (2019b) Physical fitness is predictive for 5-year survival in older adults with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities* **32**, 958–66.

Podgorski C. A., Kessler K., Cacia B., Peterson D. R. & Henderson C. M. (2004) Physical activity intervention for older adults with intellectual disability: report on a pilot project. *Mental Retardation* **42**, 272–83.

Riebe D., Ehrman J. K., Liguori G., Magal, M. & American College of Sports Medicine (2018) ACSM's guidelines for exercise testing and prescription.

Robertson S., Kremer P., Aisbett B., Tran J. & Cerin E. (2017) Consensus on measurement properties and feasibility of performance tests for the exercise and sport sciences: a Delphi study. *Sports Med Open* **3**(2), 1–10.

Shields N. & Taylor N. F. (2010) A student-led progressive resistance training program increases lower limb muscle strength in adolescents with Down syndrome: a randomised controlled trial. *Journal of Physiotherapy* **56**, 187–93.

Shields N., Taylor N. F. & Dodd K. J. (2008) Effects of a community-based progressive resistance training program on muscle performance and physical function in adults with Down syndrome: a randomized controlled trial. *Archives of Physical Medicine and Rehabilitation* **89**, 1215–20.

Surburg P. R., Suomi R. & Poppy W. K. (1992) Validity and reliability of a hand-held dynamometer applied to adults with mental retardation. *Archives of Physical Medicine and Rehabilitation* **73**, 535–9.

van der Ploeg R. J., Fidler V. & Oosterhuis H. J. (1991) Hand-held myometry: reference values. *Journal of Neurology, Neurosurgery, and Psychiatry* **54**, 244–7.

van der Ploeg R. J., Oosterhuis H. J. & Reuvekamp J. (1984) Measuring muscle strength. *Journal of Neurology* **231**, 200–3.

van Schijndel-Speet M., Evenhuis H. M., van Wijck R., van Montfort K. C. & Echteld M. A. (2016) A structured physical activity and fitness programme for older adults with intellectual disabilities: results of a cluster-randomised clinical trial. *Journal of Intellectual Disability Research* **61**, 16–29.

Vet H. C. W. D., Terwee C., Mokkink W. & Knol D. L. (2015) Measurement in medicine: a practical guide.

Weterings S., Oppewal A. & Hilgenkamp T. I. M. (2020a) The feasibility of vigorous resistance exercise training in adults with intellectual disabilities with cardiovascular disease risk factors. *Journal of Applied Research in Intellectual Disabilities* **33**, 488–95.

Weterings S., Oppewal A. & Hilgenkamp T. I. M. (2020b) Hand-held dynamometry is feasible and reliable in adults with intellectual disabilities. Under review.

Weterings S., Oppewal A., van Eeden F. M. M. & Hilgenkamp T. I. M. (2018) A resistance exercise set for a total body workout for adults with intellectual disabilities, a pilot study. *Journal of Applied Research in Intellectual Disabilities* **32**, 730–6.

Williams M. A., Haskell W. L., Ades P. A., Amsterdam E. A., Bittner V., Franklin B. A. *et al.* (2007) Resistance exercise in individuals with and without cardiovascular disease: 2007 update: a scientific statement from the American Heart Association Council on Clinical Cardiology and Council on Nutrition, Physical Activity, and Metabolism Review. *Circulation* **116**, 572–84.

WMA (2013) World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* **310**, 2191–4.

Zatsiorsky V. M., Kraemer W. J. & Fry A. C. (2021) Science and practice of strength training.

Accepted 7 April 2022