

The validity and reliability of the Functional Strength Measurement (FSM) in children with intellectual disabilities

W. F. M. Aertssen,¹  B. Steenbergen^{2,3} & B. C. M. Smits-Engelsman⁴

¹ Avans+, University for Professionals, Breda, The Netherlands

² Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

³ School of Psychology, Australian Catholic University, Melbourne, Australia

⁴ Department of Health and Rehabilitation Sciences, Division of Physiotherapy, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

Abstract

Background There is lack of valid and reliable field-based tests for assessing functional strength in young children with mild intellectual disabilities (IDs).

Aim The aim of this study was to investigate the test–retest reliability and construct validity of the Functional Strength Measurement in children with ID (FSM-ID).

Method Fifty-two children with mild ID (40 boys and 12 girls, mean age 8.48 years, SD = 1.48) were tested with the FSM. Test–retest reliability ($n = 32$) was examined by a two-way interclass correlation coefficient for agreement (ICC 2.1A). Standard error of measurement and smallest detectable change were calculated. Construct validity was determined by calculating correlations between the FSM-ID and handheld dynamometry (HHD) (convergent validity), FSM-ID, FSM-ID and subtest strength of the Bruininks-Oseretsky test of motor proficiency – second edition (BOT-2) (convergent validity) and the FSM-ID and balance subtest of the BOT-2 (discriminant validity).

Results Test–retest reliability ICC ranged 0.89–0.98. Correlation between the items of the FSM-ID and HHD ranged 0.39–0.79 and between FSM-ID and BOT-2 (strength items) 0.41–0.80. Correlation between items of the FSM-ID and BOT-2 (balance items) ranged 0.41–0.70.

Conclusion The FSM-ID showed good test–retest reliability and good convergent validity with the HHD and BOT-2 subtest strength. The correlations assessing discriminant validity were higher than expected. Poor levels of postural control and core stability in children with mild IDs may be the underlying factor of those higher correlations.

Keywords FSM, functional strength, Functional Strength Measurement, mild intellectual disabilities, reliability, validity

Introduction

Children with mild intellectual disabilities (IDs) are known to have delayed motor milestones. Pathophysiological features such as hypotonia and joint laxity are suggested to contribute to these motor problems (Dey *et al.*, 2013; Lauteslager, 2000). Motor problems make participating in sport and daily

Correspondence: Ms. Wendy Aertssen, laan 1940-1945 no. 60, 4194 RH Meteren, The Netherlands (e-mail: wendyaertssen@hotmail.com).

outdoor activities more difficult (Barr & Shields, 2011) leading to a more sedentary lifestyle, lower levels of physical activity and poor physical fitness levels (Wuang *et al.*, 2013; Lin and Wuang, 2012; Philips and Holland, 2011). Furthermore, poor physical fitness has impact on tasks of independent daily living (Gupta and Rao, 2011).

One of the components of physical fitness concerns muscle function, for example, muscle strength, muscle power and muscle endurance. Different studies reported lower levels of muscle function in children with ID in comparison with their typically developing (TD) peers (Hartman *et al.*, 2015; Golubovic *et al.*, 2012; Cowley *et al.*, 2010; Tejero-Gonzalez *et al.*, 2013; Cioni *et al.*, 1994; Horvat *et al.*, 1997; Wuang *et al.*, 2013; Mercer & Lewis 2001; Hassani *et al.*, 2014).

Various measures have been used in these studies examining different aspects of muscle function. Studies that examined strength of an isolated muscle group using handheld dynamometry (HHD) found lower levels of muscle strength in individuals with ID compared with TD peers (Cowley *et al.* 2010, Cioni *et al.* 1994, Horvat *et al.*, 1997; Wuang *et al.*, 2013, Mercer & Lewis, 2001). Studies that examined muscle power using functional strength items, such as counter jump movements (Blomqvist *et al.* 2013; Hassani *et al.*, 2014) and standing broad jump (Hartman *et al.*, 2015; Golubovic *et al.*, 2012), reported lower levels of muscle power. Yet other studies examined localised muscle endurance using sit-ups (Hartman *et al.*, 2015, Blomqvist *et al.*, 2013) and trunk extension (Blomqvist *et al.* 2013) and found lower levels in children with ID. Furthermore, using the Bruininks-Oseretsky test of motor proficiency – second edition (BOT-2), a reliable measure for children with ID (Wuang & Su, 2009), it was shown that children with ID have lower scores compared with TD children on the subtest strength, which contains both power and localised muscle endurance items (standing long jump, push-up, sit-up, wall sit and v-up).

In sum, previous studies reported lower levels of (functional) strength in individuals with ID compared with their TD peers. However, the psychometric properties of the different measures used are not well investigated. Wouters *et al.* (2017) revealed in their systematic review that some studies described the test–retest reliability, but no single study examined

children below the age of 10 years. Although the different measures were able to distinguish between TD children and those with ID (known group validity), construct validity of the all existing field-based measures is lacking. Construct validity provided information about the parameters measured, which is useful for the interpretation of test results, diagnostic purposes and the choice of intervention.

The Functional Strength Measurement (FSM) is a norm-referenced test for functional strength. It consists of eight items including muscle power (overarm and underarm throwing, standing long jump, chest pass) and muscle endurance (lateral step-up, sit to stand, lifting a box and stair climbing). The FSM is a valid and reliable measure in TD children and children with mild motor problems (Aertssen *et al.*, 2016). The items of the FSM are closely related to activities that children perform in daily life ensuring that this instrument ecologically valid. The absence of a valid and reliable instrument for functional strength in young children with mild ID raises questions about the suitability and applicability of the FSM in this population.

The aim of this study is to investigate the test–retest reliability and the construct validity of the FSM in children with mild ID aged 5–10 years. On the basis of data in TD children and children with mild motor problems, we expected good test–retest reliability for the FSM in children with mild ID [interclass correlation coefficient (ICC) > 0.8].

As there is no valid field-based outcome measure for functional strength, we investigated the construct validity by formulating three hypotheses.

- 1 To test if the FSM measures strength, we investigated the relation between the FSM and the HHD. Because isometric strength outcomes and functional strength are not linearly related (Beenakker *et al.*, 2001), we expected moderate correlations between items of the FSM and the HHD (0.4–0.7) (convergent validity).
- 2 To further test if the FSM measures strength, we also investigated the relation between the FSM and BOT-2 strength items. The BOT-2 and the FSM both include functional strength items but are different in many aspects. In the BOT-2 (subtest strength), there is more emphasis on core muscles (sit-up, push-up, v-up). Two items use

static muscle contraction (wall sit, v-up), and only one item assesses explosive power (long jump). The FSM contains items for explosive power and muscle endurance for upper and lower extremities. Therefore, moderate correlations were expected (0.4–0.7) (convergent validity).

- 3 The items of the FSM were developed with low balance demands. To examine if this was also the case in children with mild ID, we examined the relation between the FSM and BOT-2 balance items, and expected low correlations (<0.4) (discriminant validity).

Method

Design

This was a cross-sectional study. The study was approved by the medical ethics committee (ECSW2014-3107-232), and informed consent was obtained from the parents of all children.

Participants

Children who participated were recruited from special needs schools in the Netherlands and from different private paediatric physical therapy practices. Teachers and physical therapist were asked to identify eligible children matching the inclusion and exclusion criteria. The primary inclusion criterion was mild ID (IQ 50–70). Exclusion criteria were general illness, severe heart problems or neurological disorders such as cerebral palsy. The information about IQ levels and medical conditions was extracted from the school files.

Outcome measures

Anthropometric measures

Height in centimetres was measured using a tape measure affixed to the wall. Weight in kilograms was measured using an electronic scale. Children were barefoot and wore sports clothing.

Functional Strength Measurement (Smits-Engelsman & Verhoef-Aertssen, 2012)

The FSM measures two types of muscle function: (1) the explosive power generated during a movement

(measured in centimetres) and (2) muscular endurance (number of repetitions within a 30-s time frame). The standardised warm-up protocol was used as described in the manual (Smits-Engelsman & Verhoef-Aertssen, 2012). The protocol includes practice trials (with a maximum of 5) and three rated trials. The best score of the three trials was used for analysis.

A pilot study was conducted to examine the feasibility of the FSM in 20 children with ID (not included in the present study). Following this, some adaptations in the instructions were made. The first adaptation was that the instructions were given by showing a photo of a child performing the task. The second adaptation was that when performing the item sit to stand, children were encouraged to stand up and reach the experimenter's hand so it was easier to straighten up the legs. During lateral step-up task, some children wanted to walk up the stairs. Therefore, the third adaptation was that the tester sat on top of the stairs so the children were constrained to step up sideways instead of walking up and down the stairs. In the items overarm throwing, underarm throwing and standing long jump, it was difficult for the children to take the right position behind the starting line. Therefore, the fourth adaptation was that visual cues were provided consisting of a thick line with two footsteps positioned behind the line. The FSM items in TD children and the adaptations for children with mild ID are described in Fig. 1. In the present study, the FSM with adaptations for children with IDs was used (FSM-ID).

Strength and balance items of the Bruininks-Oseretsky test of motor proficiency – second edition (Bruininks & Bruininks, 2005)

The BOT-2 is a test battery for gross and fine motor skills, balance, strength, running speed and agility in children aged 4–21 years. The BOT-2 has good test-retest reliability in children with ID (ICC 0.90–0.97) (Wuang & Su, 2009). A Rasch analysis was performed with the items of the BOT-2 (Wuang *et al.*, 2009), but relationship with other measures was not investigated. We used the subset balance (nine items) and subtest strength (five items) in our study. The raw scores of these different items have to be transformed into point scores by using the conversion chart provided

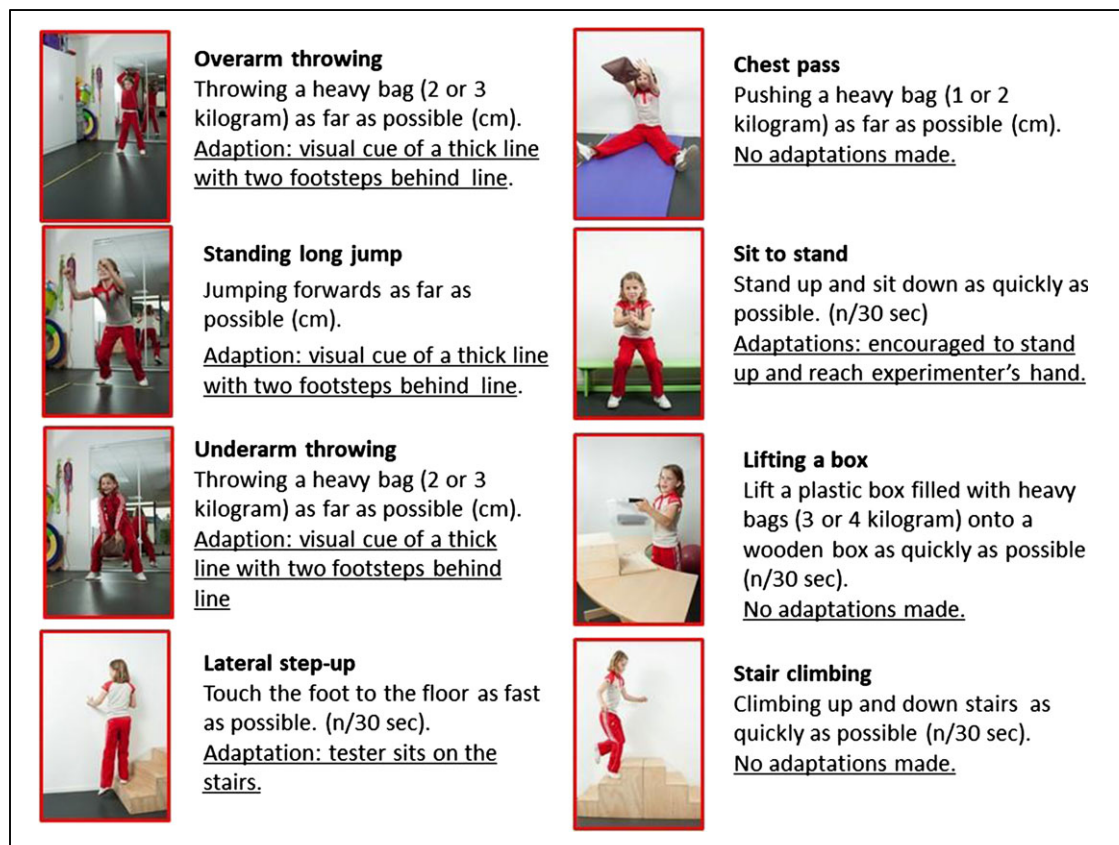


Figure 1 Items of the Functional Strength Measurement and the adaptations made for children with intellectual disabilities. cm=centimeter; n/30 sec=number of repetitions in 30 seconds [Colour figure can be viewed at wileyonlinelibrary.com]

on the score form for each item. These point scores were used in our analyses.

Handheld dynamometer

In this study, we used the MicroFET-2 (Hoggan Health Industries Inc., West Jordan, UT, USA) handheld dynamometer (HHD). The HHD is a device that measures isometric strength while a participant pushes against a portable power-transducer. There are two methods of testing, the 'make' and 'break' methods. In the make method, the participant pushes for 3 s against the HHD; the maximum force is recorded in Newton. In the break method, the tester overpowers the force-generating capacity of the participant, and the force is recorded when there arises a movement. In the present study, the break method was used according to the protocol of Beenakker (2001). We measured the flexion and extension of the

elbow and the extension of the knee. The child performed three trials, and the best score (in newton) was used in the statistical analyses. Wuang *et al.* (2013) showed that the HHD is a reliable measure for the lower extremities in children with ID (ICC 0.69–0.95, knee extension 0.93).

Procedure

First, the children were measured for body weight and height. Next, all children were tested on the FSM-ID by six trained paediatric physical therapy students. Testing occurred at the child's school, private physical therapy practice or home. Children were asked to wear sports clothing and shoes.

For the test-retest reliability study, a convenience subsample of 32 children (out of the original 52 children) was tested for a second time on the FSM-ID within a period of 2 weeks.

For the validity study, a subsample of 29 children (out of the original 52 children) was tested with the BOT-2 (subtest balance and subtest strength) and the HHD. These tests were performed 1 week after the first test day. Between the administration of the BOT-2 and HHD, there was a 15-min rest period.

Statistical analyses

The Shapiro–Wilk test showed normal distribution of the data. Body mass index (BMI) was calculated by dividing weight by the square of height [weight (kg)/height² (metre)]. To classify children with low weight, normal weight, overweight or obesity, previously reported cut-off point for children were used (Talma *et al.*, 2010).

To investigate the test–retest reliability, a two-way ICC for agreement (ICC 2.1A) was calculated using the raw scores (centimetres and the number of repetitions in 30 s) of the FSM-ID. The standard error of measurement (SEM) was calculated by dividing the standard deviation of the differences by the square root of two ($SD_{\text{difference}}/\sqrt{2}$). We also calculated the smallest detectable change (SDC 95%) by multiplying the standard deviation of the difference with 1.96 ($1.96 \times SD_{\text{difference}}$) (de Vet *et al.*, 2011).

For the convergent validity, Pearson correlations were calculated between the items of the FSM-ID and the items of the BOT-2 subtest strength (point scores) and between the FSM-ID and the HHD. Children with overweight or obesity are known to have higher levels of isometric strength of the lower extremities in non-weight-bearing positions (Takken, 2005). Therefore, partial correlations (controlling for BMI)

were performed for the correlations between FSM-ID and HHD of the lower extremities.

Discriminant validity was determined by calculating Pearson correlations between the items of the FSM-ID and the items of the BOT-2 subtest balance.

All statistical analyses were performed with SPSS, version 22 (IBM Corp, Armonk, NY, USA). α was set at 0.05.

Results

Group characteristics

In this study, 52 children with mild ID aged 5–10 years (40 boys and 12 girls, mean age 8.48 years, SD = 1.48) participated. Of these 52 children, eight children had Down syndrome, six children had other syndromes (fragile X syndrome and different deletion syndromes) and 38 children had mild ID without known co-morbidities. BMI scores showed that 20 of the 52 (38%) children who participated in this study were overweight (19%) or obese (19%). Table 1 shows the participants' demographics.

Test–retest reliability

In the test–retest part of this study, 32 children participated (see Table 1). For the item lateral step-up, the results of 24 children were used for analyses. For the item sit to stand, the results of 30 children were available because two children were not able to perform the items in a right manner. For the other items, the results of all 32 children were used for analyses. The test–retest reliability of the different items of the FSM-ID showed ICC ranged

Table 1 Participants' demographics per subsample

	Total sample (N = 52)		Test–retest study (N = 32)		Validity study (N = 29)	
	Mean (SD)	Min–max	Mean (SD)	Min–max	Mean (SD)	Min–max
Age (years)	8.48 (1.48)	5–10	8.42 (1.59)	5–10	8.69 (1.34)	7–10
Height (cm)	134.43 (12.63)	104–162	134.34 (15.13)	104–161.50	136.19 (10.02)	120–161.50
Weight (kg)	33.36 (11.64)	11.90–65.00	32.28 (12.58)	11.90–65.00	34.40 (9.96)	23–64.50
BMI (kg/m ²)	24.44 (6.77)	9.67–42.63	23.58 (6.87)	9.67–41.67	25.04 (5.99)	18.87–42.63
IQ	57.71 (6.56)	50–70	57.42 (6.94)	50–70	58.51 (5.93)	50–70

BMI, body mass index.

0.89–0.98. The SEM of the explosive power items ranged 9.29–29.32 cm, and the SDC of these items ranged 25.60–81.03 cm. The SEM of the muscle endurance items ranged 1.40–3.89 repetitions in 30 s, and the SDC of these items ranged 3.86–10.74 repetitions in 30 s.

The ICC, SEM and SDC are shown in Table 2.

Construct validity

The convergent validity with the HHD showed significant correlations for the items of the upper extremities of the FSM-ID and the flexion and extension of the elbow of the HHD in the range 0.47–0.72. Table 3 shows the correlations between the items of the upper extremities of the FSM-ID and the HHD.

For the items of the lower extremities of the FSM-ID, no significant correlations were found with the extension of the knee (HHD) except for the item stair climbing ($r = 0.39$ – 0.45). After

controlling for BMI, not only significant correlations with stair climbing ($r_{\text{bmi}} 0.43$ – 0.46) were found but also significant correlations with long jump ($r_{\text{bmi}} 0.41$) and lateral step-up left ($r_{\text{bmi}} 0.42$). The correlations with sit to stand approached significance ($r_{\text{bmi}} 0.37$ – 0.38 , $P = 0.05$ – 0.06). Table 4 shows the correlations between the items of the FSM-ID and the different muscle groups measured with the HHD.

Correlations between the items of the FSM-ID and the total point score of the subtest strength of the BOT-2 were used to test convergent validity and were significant in the range 0.41–0.80. Table 5 shows the correlations between the FSM-ID and the items of subtest strength of the BOT-2.

The discriminant validity was tested by calculating the correlations between the balance items of the BOT-2 and FSM-ID. Significant correlations yielded values above the 0.40 ($r = 0.41$ – 0.70). Table 6 shows the correlations between the FSM-ID and the items of subtest balance of the BOT-2.

Table 2 Test–retest reliability [interclass coefficient agreement (ICC)], standard error of measurement (SEM) and smallest detectable change (SDC) of the Functional Strength Measurement

N = 32	T1 mean (SD)	T2 mean (SD)	ICC (95% CI)	SEM	SDC
Overarm throwing [†]	185.78 (86.35)	189.84 (90.98)	0.89 (0.79–0.95)	29.32	81.03
Standing long jump [†]	78.79 (36.38)	75.85 (35.79)	0.93 (0.87–0.97)	9.26	25.60
Underarm throwing [†]	258.85 (113.61)	259.09 (114.78)	0.98 (0.91–0.98)	24.52	67.78
Lateral step-up right [‡]	25.43 (10.32)	25.22 (9.46)	0.92 (0.81–0.96)	2.92	8.09
Lateral step-up left [‡]	24.79 (10.45)	24.96 (10.12)	0.91 (0.79–0.96)	3.08	8.51
Chest pass [†]	174.18 (62.34)	169.27 (55.32)	0.92 (0.85–0.96)	16.46	45.49
Sit to stand [‡]	17.00 (5.93)	17.47 (6.26)	0.89 (0.78–0.94)	2.06	5.70
Lifting a box [‡]	13.00 (7.16)	13.58 (7.06)	0.96 (0.92–0.98)	1.40	3.86
Stair climbing [‡]	46.61 (17.50)	45.84 (19.78)	0.96 (0.91–0.98)	3.89	10.74

[†]In centimetres.

[‡]Number of repetition in 30 s.

Table 3 Correlations between the Functional Strength Measurement-ID and the handheld dynamometer for the upper-extremity items

N = 29 Pearson correlations	Flexion elbow right	Flexion elbow left	Extension elbow right	Extension elbow left
Overarm throwing	0.67**	0.67**	0.58**	0.47**
Underarm throwing	0.65**	0.66**	0.49**	0.48**
Chest pass	0.67**	0.65**	0.67**	0.59**
Lifting a box	0.72**	0.71**	0.60**	0.53**

* $P < 0.05$, ** $P < 0.01$.

Table 4 Correlations between the Functional Strength Measurement-ID and the handheld dynamometer for lower-extremity items

N = 29 Pearson correlations	Pearson correlations Extension knee right	Partial correlation controlling for BMI Extension knee right	Pearson correlations Extension knee left	Partial correlation controlling for BMI Extension knee left
Standing long jump	0.22 ($P = 0.26$)	0.41*	0.10 ($P = 0.61$)	0.30 ($P = 0.13$)
Lateral step-up right	0.22 ($P = 0.26$)	0.31 ($P = 0.12$)	0.20 ($P = 0.31$)	0.29 ($P = 0.14$)
Lateral step-up left	0.31 ($P = 0.11$)	0.42*	0.26 ($P = 0.17$)	0.37 ($P = 0.06$)
Sit to stand	0.31 ($P = 0.10$)	0.37 ($P = 0.06$)	0.29 ($P = 0.12$)	0.38 ($P = 0.05$)
Stair climbing	0.45*	0.46*	0.39*	0.30 ($P = 0.13$)

* $P < 0.05$, ** $P < 0.01$.**Table 5** Correlations between the Functional Strength Measurement-ID and subset strength Bruininks-Oseretsky test of motor proficiency – second edition

N = 29 Pearson	Long jump	Push-up	Sit-up	Wall sit	V-up	Total point score
Overarm throwing	0.39*	0.62**	0.47*	0.36 $P = 0.05$	0.27 $P = 0.15$	0.59**
Standing long jump	0.89**	0.44*	0.59**	0.40*	0.34 $P = 0.07$	0.80**
Underarm throwing	0.61**	0.58**	0.59**	0.40 $P = 0.06$	0.24 $P = 0.20$	0.69**
Lateral step-up right	0.47*	0.53**	0.33 $P = 0.08$	0.48**	0.35 $P = 0.07$	0.63**
Lateral step-up left	0.58**	0.58**	0.37 $P = 0.05$	0.50**	0.31 $P = 0.11$	0.70**
Chest pass	0.29 $P = 0.13$	0.49**	0.38**	0.18 $P = 0.34$	-0.06 $P = 0.78$	0.41*
Sit to stand	0.34 $P = 0.07$	0.65**	0.48**	0.49**	0.12 $P = 0.52$	0.57**
Lifting a box	0.35 $P = 0.06$	0.69**	0.48**	0.40*	0.15 $P = 0.45$	0.58**
Stair climbing	0.65**	0.59**	0.61**	0.51**	0.25 $P = 0.19$	0.75**

* $P < 0.05$, ** $P < 0.01$.

Discussion

The aim of this study was to examine the test–retest reliability and construct validity of the adapted FSM (FSM-ID) in children with mild IDs. We found a good reliability, and regarding the construct validity, two of the three hypotheses were confirmed by our study.

Test–retest reliability

In the FSM-ID (and in the original FSM), children have up to five practice trials to ascertain that they understand how to perform the item. Mercer and Lewis (2001) mentioned that practice trials help children with ID to make the right motor performance. This study showed good test–retest reliability (ICC 0.89–0.98) when using the FSM-ID protocol.

The FSM-ID scores of the children with mild ID were stable. However, eight children were unable to perform the lateral step-up (the first time and also the second time). These children did not fully straighten the tested leg or they pushed off from the floor with the other foot. Despite the fact that we ensured that children understood the task, they were only able to make the movement a few times in the correct manner and could not sustain the movement for 30 s. For some children with ID, it appeared that they did not have notion of time (cognition), so they stopped before the 30 s were over. Also, some children were not motivated to give their best performance during the 30 s (Halle *et al.*, 1999). For most children, however, this item was just difficult from a muscle endurance perspective. They started doing the item well but could not keep up the task during the 30 s. For these same reasons, two of these eight children

Table 6 Correlations between the FSM and the balance subtest of the Bruininks-Oseretsky test of motor proficiency – second edition (point scores)

N = 29 Pearson correlations	Walking on a line		Standing on one leg		Standing on line eyes closed		Walking heel to toe on a line		Standing on one leg eyes closed		Standing on one leg on balance beam		Heel to toe standing balance beam		Standing on one leg on balance beam eyes closed		Total point score
	Standing on a line	Walking on a line	Standing on one leg	Standing on one leg	Standing on line eyes closed	Walking heel to toe on a line	Walking heel to toe on a line	Standing on one leg eyes closed	Standing on one leg eyes closed	Standing on one leg on balance beam	Standing on one leg on balance beam	Heel to toe standing balance beam	Standing on one leg on balance beam eyes closed	Standing on one leg on balance beam eyes closed			
Overarm throwing	-0.54 P = 0.78	0.19 P = 0.32	0.34 P = 0.07	0.48**	0.56**	0.51**	0.54**	0.62**	0.49**	0.67**							
Standing long jump	0.25 P = 0.19	0.12 P = 0.53	0.29 P = 0.13	0.31 P = 0.10	0.42**	0.36 P = 0.05	0.60**	0.42**	0.55**	0.57**							
Underarm throwing	0.16 P = 0.42	0.30 P = 0.11	0.56**	0.52**	0.42**	0.55**	0.78**	0.50**	0.60**	0.74**							
Lateral step-up right	0.42*	-0.25 P = 0.21	0.47*	0.05 P = 0.82	0.23 P = 0.22	0.36 P = 0.06	0.31 P = 0.11	0.22 P = 0.26	0.51**	0.41**							
Lateral step-up left	0.33 P = 0.09	-0.21 P = 0.29	0.46*	0.11 P = 0.57	0.26 P = 0.19	0.46*	0.38*	0.39*	0.50**	0.49**							
Chest pass	-0.03 P = 0.99	0.10 P = 0.59	0.38*	0.48**	0.47*	0.60**	0.51**	0.55**	0.46*	0.65**							
Sit to stand	0.37 P = 0.05	0.01 P = 0.95	0.27 P = 0.16	0.11 P = 0.58	0.37*	0.44*	0.29 P = 0.13	0.33 P = 0.08	0.73**	0.48**							
Lifting a box	0.29 P = 0.13	0.02 P = 0.91	0.44*	0.31 P = 0.10	0.51**	0.58**	0.57**	0.57**	0.56**	0.68**							
Stair-climbing	0.31 P = 0.09	0.09 P = 0.65	0.37*	0.42*	0.52**	0.60**	0.53**	0.59**	0.55**	0.70**							

*P < 0.05, **P < 0.01.

were also not able to perform properly the item sit to stand.

Regarding the SEM, we found small systematic errors (below 15% of the mean) for all items. Looking at the SDC, we found some large values. For example, for the item overarm throwing, the SDC was 81 cm, meaning children would have to improve their performance by 81 cm or more before individual improvement can be concluded. These higher values are due to the variability in performance.

Construct validity

The construct validity was examined by testing three hypotheses.

Regarding the convergent validity, we found moderate to strong significant correlations (0.47–0.72) between the upper-extremity items of the FSM-ID and the strength of the flexors and extensors of the elbow as assessed with the HHD. These correlations were expected and comparable with correlations found in TD children (0.54–0.74). For the lower-extremity items of the FSM-ID, only stair climbing had a significant correlation with the extension of the knee (HHD) (0.39–0.45). This was different from the study in TD children where all lower-extremity items of the FSM showed significant correlations with the extension of the knee (0.42–0.69). In the present study, 38% of the children were overweight or obese. Because of the known relationship between isometric strength of the lower extremities and activities where children have to carry their own weight, we performed partial correlations controlling for BMI. This led to three more significant correlations between isometric strength outcomes and lower-extremity FSM items and three that approached significance.

Another possible explanation for the different correlations in comparison with those of the TD children could be related to the content of the tasks. Sit to stand is an item where, beside anaerobic muscle endurance, also postural control is needed (Whitney *et al.*, 2005). This repetitive item requires fast anticipatory postural adaptations. Every time you stand up and sit down from a chair, the centre of pressure must shift in an antero-posterior direction. Children with mild ID are known to have difficulties with postural reactions (Wuang *et al.* 2012). These fast repetitive contractions are also needed in the

one-leg stance lateral step-up item, for which we found a correlation of 0.48 with the total point score balance of the BOT-2.

A third explanation for the discrepancy could be that some children did not understand how to give their best performance or had lack of motivation especially in the muscle endurance items.

Regarding the convergent validity, significantly moderate to high correlations (0.41–0.80) were found between items of the FSM-ID and subset strength of the BOT-2. This is in line with our hypotheses. Concerning the correlations between the different items of the BOT-2 and the FSM-ID, it is evident that the item v-up (BOT-2) (static core contraction) showed no significant correlation with any of the items of the FSM-ID. The FSM-ID does not contain items with prolonged static muscle contractions. However, the wall-sit item (static contraction of the leg muscles) (BOT-2) showed significant correlations with the lower-extremity items of the FSM-ID (0.40–0.51). In wall sit, the strength of the lower extremities is tested, and although this involves static contractions, the correlation found with the FSM-ID is plausible. To perform wall sit (BOT-2) and standing long jump, lateral step-up, sit to stand, stairs climbing and lifting box (FSM), strength of the knee extensors is needed. The push-up item (BOT-2) had significant correlations with all items of the FSM-ID (0.44–0.69) and the sit-up item (BOT-2) with all items except for lateral step-up (FSM-ID) (0.38–0.61). In the items push-ups and sit-ups, localised core muscle strength plays a major role. In the FSM-ID, localised core items are not present, although the strength of core muscles is needed in all items.

The discriminant validity was determined by calculating correlations between the FSM-ID and the items of subtest balance of the BOT-2. We found moderate correlations (0.41–0.70), which were higher than we had hypothesised. In the validity study in TD children, low correlations (<0.39) were found between the FSM and the Movement Assessment for Children-second edition (MABC-2) (Aertssen *et al.*, 2016). A tentative explanation could be that another mediating factor is present in both measures. To explore this, we post hoc calculated the correlations between the items of subtest balance and subtest strength of the BOT-2. The total point score of subtest strength and the total point score of the subtest balance showed a correlation of 0.66. During

the balance items of the BOT-2, children must keep their hands on their hips. Therefore, arms cannot be used for balance, making the contribution of core stability more important in the performance of these tasks. Interestingly, the strength items push-up and sit-up had correlations of 0.59 and 0.63, respectively, with the total point score of the subtest balance. So core stability, and not just balance, might be the moderating factor between the subtest balance of the BOT-2 and the FSM-ID. Lastly, overall reduced postural control could lead to lower scores in both balance and functional strength items. Even though the balance requirements for the FSM are low for TD children, the level needed in the items may be closer to the limits of children with mild ID. For instance, Wuang *et al.* (2012) found a difference in centre of pressure displacements between TD children and children with ID when throwing a ball. Thus, decreased throwing distance in the three throwing items of the FSM-ID may have been partly caused by difficulties in postural control. Moreover, all the repetitive items (sit to stand, lifting a box, lateral step-up) require fast anticipatory postural adaptations in stance; as these reactions are slightly delayed in children with ID, this could also be the reason for the found correlation with balance items (Wuang *et al.* 2012).

Limitations

This study has some limitations. The HHD and the BOT-2 are both reliable measures in children with ID; however, the validity of these measures is unknown in children with mild ID. Validity gives us information about the construct measured. For example for the BOT-2 subset strength, we found moderate correlation with the subset balance, meaning that there is also another mediating component. As mentioned before, validity is unknown for all field-based muscle strength measures in young children with ID (Wouters *et al.* 2017). In this study, a heterogenic group of children with mild ID participated. This makes the results more generalisable but could also be of influence on the correlations found. Children with Down syndrome ($n = 8$) are smaller and are known to have more pathophysiological and health problems than do other children with ID. Children with Down syndrome often experience joint hypermobility and muscle

hypotonia (Lauteslager, 2000; Dey *et al.*, 2013), which may influence motor performance and physical fitness. It would therefore be interesting to investigate in future studies the influence of hypotonia and hypermobility on functional strength in children with mild ID.

Conclusion

Children with mild ID are known to have motor problems and lower levels of physical fitness. It is important that a reliable and valid instrument can be used to investigate the motor problems so that interventions can be based on the specific problems of the child. Therefore, we examined the test–retest reliability and construct validity of the FSM-ID in young children with mild IDs (IQ 50–70). Results showed good test–retest reliability (ICC 0.89–0.98) and good convergent validity with the HHD and BOT-2 subtest strength. The correlations between the FSM-ID and the BOT-2 (discriminant validity) were higher than expected. It is likely that poor postural control and lower core muscle strength mediate those higher correlations.

Clinical utility

The FSM-ID can be used to measure functional strength in children with mild ID. This is the first measure for functional strength (muscle power and muscle endurance in upper and lower extremities) where reliability and validity were investigated in this group of children. It is important to realise that components such as postural control and core stability may be important factors in the level of functional strength that children can produce.

Acknowledgements

We would like to thank all children and parents for participating in this study. And we thank the students of Avans+ (Annet Rijneveld, Henny Bullens, Ilse Mahieu, Wendy Tiemessen, Hilde Delisse and Esther van de Wiel) for helping with the data collection.

References

Aertssen W. F. M., Ferguson G. D. & Smits-Engelsman B. C. M. (2016) Reliability, structural and construct

validity of the Functional Strength Measurement (FSM) in children aged 4–10 years. *Physical Therapy* **96**, 888–97.

Barr M. & Shields N. (2011) Barriers and facilitators to physical activity in children with Down syndrome. *Journal of Intellectual Disability Research* **55**, 1020–33.

Beenakker E. A. C., van der Hoeven J. H., Fock J. M. & Maurits N. M. (2001) Reference values of maximum isometric muscle force obtained in 270 children aged 4–16 years by hand-held dynamometry. *Neuromuscular Disorders* **11**, 441–6.

Bloomqvist S., Olsson J., Wallin L., Wester A. & Rehn B. (2013) Adolescents with intellectual disability have reduced postural balance and muscle performance in trunk and lower limbs compared to peers without intellectual disability. *Research in Developmental Disabilities* **34**, 198–206.

Bruininks R. H. & Bruininks B. D. (2005) *Bruininks-Oseretsky Test of Motor Proficiency*, 2nd edn. Pearson Assessment, Minneapolis.

Cioni M., Cocilovo A., Di Pasquale F., Araujo M. B., Siqueria C. R. & Bianco M. (1994) Strength deficit of knee extensor muscles of individuals with Down syndrome from childhood to adolescence. *American Journal of Mental Retardation* **99**, 166–74.

Cowley P. M., Ploutz-Snyder L. L. & Baynard T. (2010) Physical fitness predicts functional tasks in individuals with Down syndrome. *Medicine and Science in Sports and Exercise* **42**, 388–93.

Dey A., Bhowmik K., Chatterjee A., Chakrabarty P. B., Sinha S. & Mukhopadhyay K. (2013) Down syndrome related muscle hypotonia: association with COL6A3 functional SNP rs2270669. *Frontiers in Genetics* **4**, 1–8.

Golubovic S., Maksimovic J., Golubovic B. & Glumbic N. (2012) Effects of exercise on physical fitness in children with intellectual disabilities. *Research in Developmental Disabilities* **33**, 608–14.

Gupta S. & Rao B. K. (2011) Effect of strength and balance training in children with Down's syndrome: A randomized controlled trial. *Clinical Rehabilitation* **25**, 425–32.

Halle J. W., Gabler-Halle D. & Chung Y. B. (1999) Effects of a peer-mediated aerobic condition program on fitness levels of youth with mental retardation: two systematic replications. *Mental Retardation* **37**, 435–48.

Hartman E., Smith J., Westendorp M. & Visscher C. (2015) Development of physical fitness in children with intellectual disabilities. *Journal of Intellectual Disability Research* **59**, 439–49.

Hassani A., Kotzamanidou M. C., Tsimaras V., Lazaridis S., kotzamanidis C. & Patikas D. (2014) Differences in counter-movement jump between boys and girls with and without intellectual disability. *Research in Developmental Disabilities* **35**, 1433–8.

Horvat M., Pitetti K. H. & Croce R. (1997) Isokinetic torque, average power and flexion/extension ratios in nondisabled adults and adults with mental retardation.

- Journal of Orthopaedic and Sports Physical Therapy* **25**, 395–9.
- Lauteslager P. E. M. (2000) *Children with Down Syndrome, Motor Development and Intervention*. Universiteit Utrecht, Utrecht.
- 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.
- Mercer V. S. & Lewis C. L. (2001) Hip abductor and knee extensor muscle strength of children with and without Down syndrome. *Pediatric Physical Therapy* **13**, 18–26.
- Philips A. C. & Holland A. J. (2011) Assessment of objectively measured physical activity levels in individuals with intellectual disabilities with and without Down's syndrome. *PLoS One* **6**, 1–7.
- Smits-Engelsman B. C. M. & Verhoef-Aertssen W. F. M. (2012) *Functional Strength Measurement (FSM): Manual*. FSM production, Meteren. Available at: <http://www.functionalstrengthmeasurement-fsm.com>.
- Takken T. (2005) Richtlijnen Conditietraining Kinderen met obesitas. In: *Afd Kinderfysiotherapie & Pediatrische Inspanningsfysiologie WKZ*. UMC, Utrecht.
- Talma H. *et al.* (2010) *Groeidiagrammen 2010*. TNO, Den Haag.
- Tejero-Gonzalez C. M., Martinez-Gomez D., Bayon-Serna J., Izquierdo-Gomez R., Castro-Pinero J. & Veiga O. L. (2013) Reliability of the alpha health-related fitness test battery in adolescents with Down syndrome. *Journal of strength and conditioning* **27**, 3221–4.
- de Vet H. C. W., Terwee C. B., Mokkink L. B. & Knol D. L. (2011) *Measurement in Medicine*. Cambridge university press, New York.
- Whitney S. L., Wrisley D. M. & Marchetti G. F. (2005) Clinical measurement of sit-to-stand performance in people with balance disorders: validity of data for the Five-Times-Sit-to-Stand Test. *Physical Therapy* **85**, 1034–45.
- Wouters M., Evenhuis H. M. & Hilgenkamp T. I. M. (2017) Systematic review of field based physical fitness tests for children and adolescents with intellectual disabilities. *Research in Developmental Disabilities* **61**, 77–94.
- Wuang Y. P. & Su C. Y. (2009) Reliability and responsiveness of the Bruininks-Oseretsky Test of Motor Proficiency – Second Edition in children with intellectual disability. *Research in Developmental Disabilities* **30**, 847–55.
- Wuang Y. P., Lin Y. H. & Su C. Y. (2009) Rasch analysis of the Bruininks-Oseretsky Test of Motor Proficiency – Second Edition in intellectual disabilities. *Research in Developmental Disabilities* **30**, 1132–44.
- Wuang Y. P., Long I. M. & Liu M. F. (2012) Relationships between task-oriented postural control and motor ability in children and adolescents with Down syndrome. *Research in Developmental Disabilities* **33**, 1792–8.
- Wuang Y. P., Chang J. J., Wang M. H. & Lin H. C. (2013) Test–retest reliabilities of hand-held dynamometer for lower-limb muscle strength in intellectual disabilities. *Research in Developmental Disabilities* **34**, 2281–90.

Accepted 10 May 2018