2016-05-01

Intra-rater reliability of clinical measures of leg function, in typically developing children aged 1-4 years.

Marsden, Jonathan

http://hdl.handle.net/10026.1/4993

Association of Paediatric Chartered Physiotherapist Journal

This is an accepted article published by the ACPC Journal and available at

http://apcp.csp.org.uk/group-journal/apcp-journal-volume-7-number-1

Intra-rater reliability of clinical measures of leg function, in typically developing children aged 1-4 years.

Rachel Rapson [a*], Professor Jonathan Marsden [b]

[a] Torbay and South Devon Healthcare NHS Trust

[b] School of Health Professions, Faculty of Health, Plymouth University

**Corresponding author:* rachel.rapson@nhs.net

ABSTRACT

Purpose

To develop a battery of measures of leg range of motion, muscle tone and indicators of strength for use in clinical trials in non-ambulant children with cerebral palsy, and to test intra-rater reliability in typically developing (TD) young children.

Method

Fifteen healthy children, five boys, with a mean age of 29.6 months (SD 9 months) were tested. Equipment including a footboard, digital inclinometers, Myotonometer, Ultrasound and tape measurement was used. The measures were repeated seven days later.

Results

The Intra-class Correlation Coefficient (ICC) reached acceptable levels of reliability for seven of the fifteen measures; slow hamstring (ICC=0.84), fast hamstring (ICC=0.79), Myoton F (ICC=0.74) and Myoton S (ICC=0.73), muscle girth (ICC=0.95), ultrasound circumference (ICC=0.71) and depth (ICC=0.76) of rectus femoris.

Conclusion: The battery was feasible for use in the home with young children. Intra-rater reliability was shown for seven tests for TD children.

Introduction

Clinical trials aimed at evaluating interventions with children with cerebral palsy (CP), require outcome measures that are reliable over long periods of time. Young children pose several problems in the generation of a reliable test battery to measure lower limb function. The ability to follow instructions at this age is limited especially in the presence of additional cognitive deficits. This makes it difficult to test common clinical outcomes such as maximal voluntary muscle contraction that rely on following complex commands. A child’s attention span can be short, meaning that tests need to be engaging to allow for standardisation of position and degree of background muscle activity. Finally, children with severe developmental delay may fatigue easily meaning that overall battery duration should be short and position changes between tests should be minimised.

Studies of normal ranges of movement (ROM) in typically developing (TD) children highlight that range of movement changes with age (Kilgour et al., 2002, Soucie et al., 2011), particularly between birth and five years. There is minimal data on the reliability of measures of ROM in TD young children. Traditional goniometry has been shown to be reliable in older non-ambulant children with CP (Fosang et al., 2003, Bartlett et al., 1985) with levels of measurement error of between 10-28˚ (Kilgour et al., 2003, Stuberg et al., 1988). The greatest variances were found when measuring bi-articular muscles (McDowell et al., 2000) where there is an increased incidence of spasticity and contracture.

Reliability can be reported using the Intra-class correlation coefficient (Shrout and Fleiss, 1979) . The use of digital inclinometers was reported in measurement of hip abduction, producing good levels (ICC >0.85) of intra-rater reliability (Herrero et al., 2011). A factor that may contribute to the low reliability when measuring range of motion is the variability in the duration and size of the applied force used to move a joint to its end range. Maas et al (2012) suggested standardising the applied force when measuring range of motion in the ankle by using a hand held dynamometer, which includes a torque wrench and goniometer attached to a footboard. Variations in the resistive torque are significant for movements about the knee and hip, due to the increased length and weight of the limb moving in relation to gravity. Therefore, for the tests of hamstring and hip flexor extensibility the application of force are manually determined in the clinical setting.

Hypertonia in children with CP is caused by both changes in passive musculo-tendinous properties and enhanced stretch reflexes (spasticity). The different components of muscle tone can be difficult to assess (Pandyan et al., 1999). Clinical tests such as the Tardieu scale aim to differentiate between the components by scoring the resistance to movements above and below the stretch reflex threshold (Scholtes et al., 2006, Boyd and Graham, 1999, Gracies et al., 2000). More recently, ultrasound has been used to determine factors such as muscle fascicle length and stiffness in the presence of spasticity(Kwah et al., 2013). These techniques require access to computer controlled motors (e.g. dynamometry) and make the measure unfeasible for clinical trials when measures may be taken outside of the laboratory setting. The Myotonometer is a portable device that measures tissue compliance. It applies a standardised perturbation via a probe and the subsequent motion of the probe is determined using an accelerometer; from this a range of measures of tissue compliance are provided such as tissue stiffness and creep (Lidstrom et al., 2009). It is a highly reliable tool in healthy adult subjects (ICC> 0.84) (Leonard et al., 2003). In children with CP over four years old the test-retest reliability was substantial in the relaxed medial gastrocnemius (ICC>0.89)(Aarrestad et al., 2004). It is portable and feasible to use in this population.

Muscle cross-sectional area or thickness can be measured by ultrasound and may be a useful quantitative measure when evaluating strengthening interventions in children who cannot comply with conventional strength tests (Ohata et al., 2008). Muscle thickness, of rectus femoris and vastus lateralis were shown to be a good predictor of muscle strength in older non-ambulant children with CP (Moreau et al., 2010). These tests provide a useful indicator of muscle strength; however reliability of these tests in TD young children has not been assessed to date.

A protocol was developed in line with the literature to assess lower limb function; range of movement, muscle tone and strength.

The battery of tests were trialled with a group of TD children aged 1-4 years to assess the ease of application and to test the intra-rater reliability for this group. Intra-rater reliability only was of interest as the battery of measures was to be used by a single rater in a subsequent study.

Method

Ethical approval was provided via the UK South West NHS Research Ethics Committees and the Faculty of Health and Human sciences, Plymouth University.

A power calculation indicated that thirteen children would be needed to demonstrate an ICC of > 0.7(power=0.85; α=0.05). This is in line with other reliability studies in this population(Aarrestad et al., 2004, McDowell et al., 2000).

Children were recruited via adverts at local nurseries and play centres. Children were included if they did not have any orthopaedic or neurological symptoms that could affect lower limb movement. They were excluded if they showed signs of infection and illness lasting more than one day in the week before the study or in the inter-measurement period. Children participated following the informed written consent of the guardians. The child was first familiarised with the tests by demonstrating them on a teddy bear and assent gained where possible.

Development of the battery of tests

The battery of nine tests was developed giving fifteen outcome measures. A Physiotherapist with fifteen years of paediatric experience in clinical examinations, such as range of motion, was trained in using the tests. The rater had no prior experience using the equipment; therefore a training period was undertaken for 4 weeks prior to beginning the trial.

The tests were carried out in the families' home using portable equipment. The testing protocol is summarised in Table 1. One leg was tested as results from both legs may well have been statistically similar, therefore only the left leg was tested for each child. The child was made comfortable with a small pillow and the parent was encouraged to comfort the child and help to stabilise the starting position. When prone, the head was turned to the same side as the tested leg to standardise influence on lower limb stretch reflexes(Aiello et al., 1992). At the beginning of the session a tape measure determined the distance between bony points on the shank and thigh and skin markers were used to indicate the measurement points for the ultrasound and myotonometer using an eye liner pencil. A compliance score was given using a simple four point scale. This was used to be able to compare whether the child's performance was similar between tests or between the two testing sessions as changes in test compliance may particularly be a factor affecting test reliability in this age group. The tests were repeated a week later by the same rater, at the same time of day and in the same setting.

Measurement of range of movement

Each range of movement test involved a warm up of three movements and three test movements were carried out with a metronome (1Hz tone) to pace the speed of the test so that movements were performed at ~5^{o}/s. The starting position, equipment and test movement for each measurement are shown in **Table 1**.

	Measurement	Start position	Equipment	Test Movement	Outcome measure
1	Hamstring -slow	Supine: tested hip flexed to 90° with knee flexed, other leg extended and resting on bed	Dual inclinometers mounted on shin and thigh pads	Extend knee through range at 5°per second. Measure popliteal angle between lower leg and line to vertical.	Passive range of motion of hamstrings
2	Hamstring- fast			Extend knee through range at 100°per second Measure popliteal angle between lower leg and line to vertical.	Muscle tone in hamstrings
3	Hip extension- slow	Prone: Staheli test- Pelvis stabilized by hand, both hips flexed over end of bed	Single inclinometer on thigh pad	Extend hip through range at 5°per second	Passive range of motion of hip flexors
4	Hip extension- fast			Extend hip through range at 100°per second	Muscle tone in hip flexors
5	Ankle dorsiflexion- slow	Prone: Knee extended with ankle over end of the bed in plantar flexion, shank resting parallel to bed	Footboard with single inclinometer and torque wrench to standardize applied force.	Dorsiflex ankle through range at 5°per second until 2Nm force reached	Passive range of motion Triceps surae
6	Ankle dorsiflexion- fast			Dorsiflex ankle through range at 100°per second until 2Nm force reached	Muscle tone triceps surae
7	Myotonometer medial head gastrocnemius	Prone: Knee extended with leg resting on bed. Test point marked at one third of the distance between the lateral epicondyle and the tip of the lateral malleolus and medial to the midline of the leg	Myoton Pro®	The machine was held vertically over the measurement point and a 0.6 N tap was applied (15 ms duration)	F=Frequency Hz D=logarithmic decrement S=Dynamic stiffness N/m C= Creep ratio R=mechanical stress relaxation time
8	Thigh circumference	Long sitting- test point marked , one third of the distance from the tip of the patella to the ASIS, in the midline of the thigh	Tape measure	Tape was run around thigh and measure read	Muscle girth Rectus Femoris
9	Ultrasound Circumference and Cross-sectional area Rectus Femoris		Ultrasound	Horizontal to direction of Rectus Femoris. Depth set half way between skin and bone.	Cross-sectional area Rectus Femoris Circumference Rectus Femoris
	Ultrasound Rectus Femoris Depth			Longitudinal to Rectus Femoris. Depth set half way between skin and bone.	Muscle depth Rectus Femoris

Table 1- Outcome Measure; Starting Position, Test and Equipment.

Ankle range of dorsiflexion with knee extension (**Figure 1** top left) was measured using a footboard (James Leckey Designs) to ensure neutral position of the ankle and midfoot. It was mounted with a digital inclinometer (Acumar®) and a digital torque wrench (Topeak®) placed as close to the ankle axis as possible on the side of the footboard. The digital torque wrench emitted an auditory tone when the desired torque (2Nm) was reached and the angle was then read on the digital display of the inclinometer.

Figure 1- Outcome Measures using the Footboard, Digital Inclinometers and Myotonometer.

Hip flexor extensibility was measured using the Staheli prone angle test (Bartlett et al., 1985)(**Figure 1**, top right) using a thigh support mounted with a digital inclinometer. The position of the proximal thigh and lumbar segments were standardised by allowing the opposite leg to hang over the side of the bed and fixation applied to the pelvis. Hamstring extensibility was determined by measuring the popliteal angle, using a thigh and shin support mounted with dual- digital inclinometers with the hip angle supported at 90° flexion, by visual estimation (Figure 1, bottom left).

Measurement of muscle tone

The Tardieu scale was used to rate the hip flexors, hamstrings and triceps surae using the slow range of movement 5°/s as V1 and repeating the range of movement (without the torque wrench) at a speed of 100°/s as V2, paced using the metronome. Three warm up tests and three measures were taken using the digital inclinometers as before.

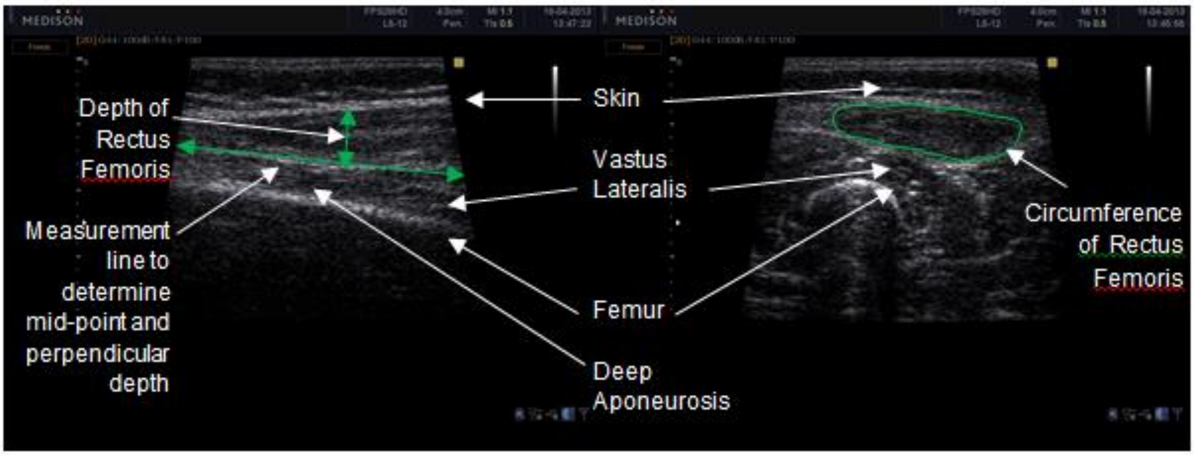
The myotonometer (MyotonPRO®, London, UK) was used to record muscle tone in the medial head of the left gastrocnemius. (Figure 1, bottom right). A small 0.6 N tap (15 ms duration) was applied and the mechanical response recorded via the accelerometer housed within the MyotonPRO®. The oscillation frequency of the muscle response, the decrement in the oscillation over time and the visco-elastic stiffness of the tissue was recorded.

Measurement of muscle size

The participant was positioned in long- sitting on an examination couch with both legs lying flat. A portable ultrasound machine (MySono U5), with an L5-12 linear transducer, was used to measure the cross-sectional area and thickness of rectus femoris, as an indicator of strength. The level of ultrasound used was below the levels set by The British Medical Ultrasound Society(Hoskins et al., 2010).The ultrasound head was horizontally aligned with the thigh marker and the focus was set to half way between the skin and the bone. Three images of the thigh were taken with the participant at rest and the average taken. The ultrasound head was then turned through 90˚ and aligned longitudinally at the same point and another three images were taken.

The measurements were made using the calliper function on the machine to draw the circumference of the muscle on the image. The machine then calculated the cross-sectional area. The calliper function was used on the transverse image to determine the midpoint of the image and to draw a vertical line to measure the depth of the muscle at that midpoint (**Figure 2**).

Figure 2- Ultrasound Images of Rectus Femoris Depth and Circumference



Thigh circumference was measured at the same point using an anatomical measuring tape. This was repeated three times with the tape removed between measures. A simple four point compliance score was given for each child, in each test session, rating compliance from poor to excellent.

Analysis

The average of the three data points was taken for each measure taken and the mean score and standard deviation calculated. Normal distribution of the data was tested using the Shapiro-Wilks test.
The outcome measures generated normally distributed ratio data which were explored using the Intra-class Correlation Coefficient (ICC) to determine the relative reliability of the measures. The 2:1 ICC model was used- a two-way random effects single measures model of absolute agreement (Shrout and Fleiss, 1979) . Where the ICC approaches +1 the correlation is said to be strong, indicating better reliability. The ICC should be >0.7 for the measure to be acceptable for use in subsequent clinical trials (Fitzpatrick et al., 1999). The Standard Error of Measurement (SEM2) was calculated to provide an absolute index of reliability and an estimate of the precision of the scores. In line with methods described by Weir (2005) this was defined as: SEM2 = SD √ (1- ICC).
Intra-rater reliability was further assessed using Bland Altman plots to determine limits of agreement. These provide a visual display to compare the variability of each pair of measures using the mean and two standard deviations and the detection of systematic error (Bland and Altman, 1999).

Results

Fifteen children (aged 29.6 ± 9.0 months, 5 male: 10 female) were recruited.
The youngest participants were both 17months old and needed breaks during testing. One of these children declined the range of movement tests and the other declined testing of the ankle range of movement. The battery of tests took 40 minutes to complete, including a demonstration in the first session, and took 30 minutes in the follow up session.

The Shapiro-Wilks test demonstrated normal distribution of the data. Table 2 shows the mean score and standard deviation alongside the ICC. The mean range of movement concurs with that found in the literature, with the exception the hip extension score which was significantly higher than previously reported (Kilgour et al., 2002).

The reliability testing of the fifteen outcome measures demonstrated that seven measures reached acceptable levels of reliability greater than 0.7 (Fitzpatrick et al., 1998). Dual inclinometers produced good reliable results (ICC=0.84-0.87) whereas the use of single inclinometers showed poor reliability (ICC= -0.07-0.29). The hip extension measure showed no correlation between test and re-test measures (ICC=-0.07).
The myotonometry results showed that the frequency of the muscle response (Myoton F) and the stiffness (Myoton S) were reliable.

Outcome Measure	N	Mean Score	Standard Deviation	ICC	SEM2
Hamstring –Slow (°)	14	19.13	11.88	0.84	12.61
Hamstring –Fast (°)	14	22.45	10.91	0.79	11.83
Hip extension – Slow (°)	14	27.1	4.8	0.07	_
Hip extension –Fast (°)	14	28.18	4.48	0.24	_
Ankle dorsiflexion –Slow (°)	13	25.24	4.25	0.24	_
Ankle dorsiflexion –Fast (°)	13	27	4.83	0.29	_
Myoton F -Oscillation Frequency (Hz)	15	14.72	1.12	0.74	1.25
Myoton D -Logarithmic Decrement (ratio)	15	1.07	0.11	0.21	_
Myoton S -Dynamic Stiffness (N/m)	15	234	29.04	0.73	32.28
Myoton C -Creep (ratio)	15	1.22	0.1	0.37	_
Myoton R -Mechanical Stress Relaxation Time (ms)	15	20.76	1.81	0.45	_
Thigh Muscle Girth (cm)	15	27.3	1.82	0.95	1.84
Ultrasound CSA Rectus Femoris(cm^2)	15	1.19	0.32	0.63	_
Ultrasound Circumference Rectus Femoris (cm)	15	4.62	0.71	0.71	0.81
Ultrasound Depth Rectus Femoris (cm)	15	0.81	0.14	0.76	0.15

Table 2- Intra-Class Correlation Coefficient and Standard Error of Measurement

The SEM2 was calculated for those outcomes with acceptable reliability (ICC ≥0.7).

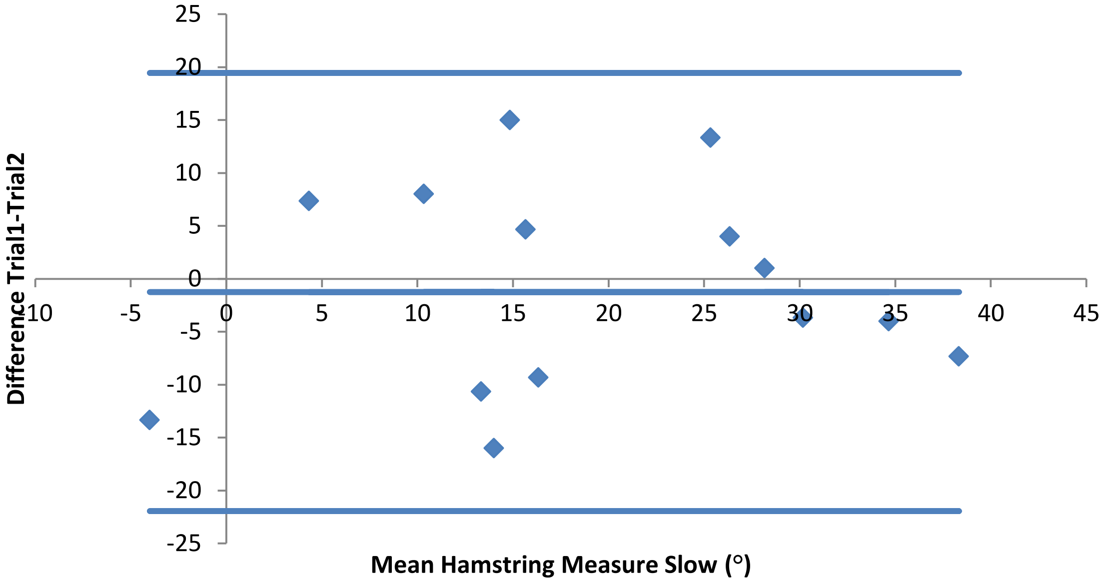


Figure 3 - Bland Altman Plot for the Slow Hamstring Range of Movement (n=14)

One of the Bland Altman plots is displayed in Figure 3. The mean difference between Trial 1 and Trial 2 is very close to zero and all the data points fall within two standard deviations of the mean. It confirms that there is no systematic error between the Trials and shows visually that there were no outlying data points. One data point below zero shows a child with a very large range who had a negative popliteal angle score.

The Bland Altman plots showed a slight positive mean difference for the fast hip extension measure, suggesting that the measures were consistently inflated by around 2° on the first trial. There were several outlying measures for the ultrasound which could be traced back to poor image quality. However, the majority of data points fell well within the two standard deviations of the mean difference of 0.

Discussion

This study developed a battery of tests of leg function and measured their reliability in a cohort of typically developing children. We found that measures of muscle tone using the myotonometer (frequency and stiffness), hamstring extensibility with slow and fast stretches, Ultrasound measures of rectus femoris circumference and depth and thigh girth were reliable (ICC>0.7).There were some potential issues affecting outcome measure reliability in this age group.

We found a noticeable variability in the compliance and positioning of the child. This sometimes added pressure on the rater to take measurements quickly when a child showed signs of being unsettled which may cause some rater error. The minimum age for consistently good compliance in TD children was found to be 20 months.

In the prone position, the children were often not able to keep their heads turned to one side. Head turning could have affected the stretch (H) reflex activity in the leg on the opposite side to which the head was turned (Aiello et al., 1992) making the limb less relaxed.

This study is the first to report results for the myotonometer in typically developing children aged 1-4years and they are similar to those reported in adults (Bailey et al., 2013). The Myoton S, measuring stiffness of the medial head of gastrocnemius increased with age indicating the increase in muscle stiffness with age which is seen clinically. The fast range of movement was measured as part of the Tardieu scale (V3) and in the absence of increased muscle tone the angle at the end of range was consistently larger at the faster speed than the slower. In children with CP, smaller ranges of overall movement would be expected at the fast speed due to the 'catch' of the abnormal stretch reflex in the presence of spasticity.

The study design assumed that the pelvis and shank were supported in a horizontal plane for the hip extension and ankle dorsiflexion measures. During testing it became apparent that the angle of the shank varied depending on the bulk of the muscles on the anterior of the tibia. The pelvis may not have remained horizontal during the hip extension measure and may have anteriorly tilted at the end of range, causing a larger angle to be consistently recorded. This resulted in poor reliability for both measures. The digital inclinometer measures exact angles to two decimal places in relation to either a horizontal or vertical axis, whereas the traditional goniometer measures the angle between two arms using a visual scale, often rounded to the nearest 5˚. The hamstring measure included dual inclinometers where the exact angle was measured between the two devices, producing good accuracy and reliability. The positioning for further reliability of ankle dorsiflexion and hip extension testing should be changed to allow a second inclinometer to be used.

The measurements of the hamstrings and ankle dorsiflexion did not measure the exact anatomical axes. When measures were taken with the inclinometers, the axes were along the shin pads and the footplate. The ankle measure used a fixed fulcrum on the footplate with the actual anatomical fulcrum varying slightly between participants. These factors need to be considered as they affect the construct validity of the measurement device as it could be argued that while they are more accurate, they are not measuring the true joint range.

At the outset of the study an attempt was made to apply a standardised torque. One limitation to this method was the difficulty in applying the torque at a standard distance from the fulcrum due to varying leg lengths. Secondly, the weight of the limb varied for each child and changed throughout range with the effect of gravity and an opposing torque lessening as the limb approached the vertical. It may be possible to use an on-line computer generated algorithm, taking into account the length of the limb and angle in relation to gravity in order to apply a standardised torque, but this was beyond the remit of this study.

The dimensions of the leg differed greatly from the youngest to the oldest child. This was adjusted for with two different length and width elasticated straps for the footboard, as the wider strap limited dorsiflexion in the smallest ankles. It was difficult to apply an effective force at the ankle to keep the smallest feet from moving in the footboard laterally and also to keep the heel down during dorsiflexion. Ultrasound imaging was more difficult for the children with the shortest femurs. The rectus femoris muscle in the shortest children tapered across the width of the image due to a shorter overall muscle length. The measure was taken at a midpoint on the captured image at 90˚ to the deep aponeurosis. In three cases this midpoint was close to the section which steeply graduated from broad to narrow, making the measurement more variable. The ultrasound measures should more accurately represent muscle size as it is possible to distinguish between skin, subcutaneous fat and muscle, whereas the thigh girth gives an indication of the bulk of both tissues combined. Previous studies have developed equations for estimated lean muscle mass based on measures of limb circumference and skin-fold thickness (Moritani, 1979) that are correlated with computerised tomography based measures of muscle circumference (Defreitas et al, 2010). Given the high reliability of thigh circumference the addition of skinfold measurements, to estimate lean muscle bulk in the age group, may be warranted. However, aside from the subcutaneous fat, it should also be noted that intramuscular fat and fibrosis

could also have contributed to the ultrasound measurement taken.
A limitation of this trial was testing only one leg, which may have artificially reduced the time needed to undertake the battery if data from both legs were needed. Additionally the study only tested intra-rater reliability, which limits the ability to generalise to studies that require inter-rater reliability.
While the results of this trial do not directly translate to the population of children with CP, some issues for consideration have arisen from this trial. Non-ambulant children with CP of the same age might be more compliant as they are used to assuming these positions and being passively moved during therapy. Conversely, some children with cognitive difficulties may find engagement with some of the tests more challenging, especially where they are required to wear some of the measurement devices.

Children with CP frequently have persistent asymmetrical tonic neck reflex, causing increase in flexor muscle tone on the side to which the head is turned and increased extensor tone in the opposite side of the body. The Thomas test, undertaken in supine with head in midline, would be preferable for future battery in children with CP to control for the influence of head turning in prone and has been found elsewhere to have similar reliability to the Staheli test (Glanzman et al., 2008, Mutlu et al., 2007).

Conclusion

A battery of tests was designed to measure neuromuscular leg function in typically developing young children (1-4years). The test battery included measures of range of movement, muscle tone and indicators of muscle strength at the hip, hamstrings and ankle.

The battery of tests was shown to be feasible to carry out in the home with young children. Novel testing methods were developed and trialled in an attempt to improve accuracy in clinical measurement. In particular the use of dual digital inclinometers, and the myotonometer provided excellent reliability results and should be considered for use in future trials.

As the ultimate target group are children with developmental delay and CP who are non-ambulant, previous work investigating the reliability of outcome measures in these groups, as well as studies in TD children informed the selection of the tests used. This battery of tests will need further development to improve the reliability of those tests that have not achieved acceptable levels and reliability testing with children with CP prior to use in clinical trials.

References

Aarrestad D, Williams M.D, Fehrer S, Mikhailenok E and Leonard C.T (2004) Intra-and interrater reliabilities of the myotonometer when assessing the spastic condition of children with cerebral palsy. *Journal of child neurology,* **19,** 894-901.

Aiello I, Rosati G, Sau G, Lentinu M, Tidore B, Sotgiu S, Cacciotto R, Posadinu D, Muzzu S and Manca I (1992) Interaction of tonic labyrinth and neck reflexes in man. *The Italian Journal of Neurological Sciences,* **13,** 195-201.

Bailey L, Samuel D, Warner M and Stokes M (2013) Parameters representing muscle tone, elasticity and stiffness of biceps brachii in healthy older males: symmetry and within-session reliability using the MyotonPRO. *Journal of Neurological Disorders,* **1,** 1-7.

Bartlett M.D, Wolf L.S, Shurtleff D and Stahell L.T (1985) Hip flexion contractures: a comparison of measurement methods. *Archives Of Physical Medicine And Rehabilitation,* **66,** 620-625.

Bland J.M and Altman D.G (1999) Measuring agreement in method comparison studies. *Stat Methods Med Res,* **8,** 135-60.

Boyd R.N and Graham H.K (1999) Objective measurement of clinical findings in the use of botulinum toxin type A for the management of children with cerebral palsy. *European Journal of Neurology,* **6,** s23-s35.

Defreitas J.M, B.T, Stock M.S, Dillon M.A, Sherk V.D, Stout J.R and Cramer J.T (2010) A comparison of techniques for estimating training-induced changes in muscle cross-sectional area. *J Strength Cond Res.* , Sept 24**,** 2383-9.

Fitzpatrick R, Davey C, Buxton, M and Jones D (1998) Evaluating patient-based outcome measures for use in clinical trials: a review. *Health Technology Assessment,* **2,** 74.

Fitzpatrick R, Shortall E, Sculper M, Murray D and Morris R (1999) Primary total hip replacement surgery: a structured review of outcomes and modelling of cost-effectiveness associated with different prostheses. *Health Technology Assessment*, **2**, 64.

Fosang A.L, Galea M.P, Mccoy A.T, Reddihough D.S and Story I (2013) Measures of muscle and joint performance in the lower limb of children with cerebral palsy. *Developmental Medicine & Child Neurology,* **45,** 664-670.

Glanzman A.M, Swenson A.E and Kim H (2008) Intrarater range of motion reliability in cerebral palsy: a comparison of assessment methods. *Pediatr Phys Ther,* **20,** 369-72.

Gracies J.M, Marosszeky J.E, Renton R, Sandanam J, Gandevia S.C and Burke D (2000) Short-term effects of dynamic Lycra splints on upper limb in hemiplegic patients. *Archives of Physical Medicine and Rehabilitation,* **81,** 1547-1555.

Herrero P, Carrera P, García E, Gómez-Trullén E.M and Oliván-Blázquez B (2011) Reliability of goniometric measurements in children with cerebral palsy: a comparative analysis of universal goniometer and electronic inclinometer. A pilot study. *BMC Musculoskeletal Disorders,* **12,** 155-155.

Hoskins P.R, Martin K and Thrush A (2010) *Diagnostic ultrasound: physics and equipment*, Cambridge University Press.

Kilgour G, McNair P and Stott N.S (2003) Intrarater reliability of lower limb sagittal range-of-motion measures in children with spastic diplegia. *Developmental Medicine & Child Neurology,* **45,** 391-399.

Kilgour G.M, McNair P.J and Stott N.S (2002) Lower limb sagittal range of motion: reliability of measures and normative values. *New Zealand Journal of Physiotherapy,* **30,** 8-24.

Kwah L.K, Pinto R.Z, Diong J and Herbert R.D (2013) Reliability and validity of ultrasound measurements of muscle fascicle length and pennation in humans: a systematic review. *Journal of Applied Physiology,* **114,** 761-769.

Leonard C.T, Deshner W.P, Romo J.W, Suoja E.S, Fehrer S.C and Mikhailenok E.L (2003) Myotonometer intra- and interrater reliabilities. *Arch Phys Med Rehabil,* **84,** 928-32

Lidstrom A, Ahlsten G, Hirchfeld H and Norrlin S (2009) Intrarater and interrater reliability of Myotonometer measurements of muscle tone in children. *J Child Neurol*, **24**, 267-74.

Maas J. C, Dallmeijer A. J, Huijing P. A, Brunstrom-Hernandez J. E, Van Kampen P. J, Jaspers R. T and Becher J. G (2012) Splint: the efficacy of orthotic management in rest to prevent equinus in children with cerebral palsy, a randomised controlled trial. *BMC Pediatr*, **12,** 38.

Mcdowell B. C, Hewitt V, Nurse A, Weston T. and Baker R (2000) The variability of goniometric measurements in ambulatory children with spastic cerebral palsy. *Gait & posture*, **12**, 114-121.

Moreau N. G, Simpson K. N, Teefey S. A and Damiano D. L (2010) Muscle architecture predicts maximum strength and is related to activity levels in cerebral palsy. *Phys Ther*, **90**, 1619-30.

Moritani T. A. D., Ha (1979) Neural factors versus hypertrophy in the time course of muscle strength gain . *Am J Phys* Med **58**, 115-130.

Mutlu A, Livanelioglu A and Gunel M. K (2007) Reliability of goniometric measurements in children with spastic cerebral palsy. *Med Sci Monit*, **13**, Cr323-9.

Ohata K, Tsuboyama T, Haruta T, Ichihashi N, Kato T and Nakamura T (2008) Relation between muscle thickness, spasticity, and activity limitations in children and adolescents with cerebral palsy. *Dev Med Child Neurol*, **50**, 152-6.

Pandyan A. D, Johnson G. R, Price C. I. M, Curless R. H, Barnes M. P. and Rodgers H (1999) A review of the properties and limitations of the Ashworth and modified Ashworth Scales as measures of spasticity. *Clinical Rehabilitation*, **13**, 373-383.

Scholtes V. A, Becher J. G, Beelen A and Lankhorst G. J (2006) Clinical assessment of spasticity in children with cerebral palsy: a critical review of available instruments. *Dev Med Child Neurol*, **48**, 64-73.

Shrout Pe F. J (1979) Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin,* **86**, 420-8.

Soucie J. M, Wang C, Forsyth A, Funk S, Denny M, Roach K. E and Boone, D (2011) Range of motion measurements: reference values and a database for comparison studies. *Haemophilia*, **17**, 500-7.

Stuberg W. A, Fuchs R. H and Miedaner J. A (1988) Reliability of goniometric measurements of children with cerebral palsy. *Dev Med Child Neurol*, **30**, 657-66.

Weir J. P (2005) Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*, **19**, 231-40.