OPEN

Novel Strength Test Battery to Permit Evidence-Based Paralympic Classification

Emma M. Beckman, PhD, Peter Newcombe, PhD, Yves Vanlandewijck, PhD, Mark J. Connick, PhD, and Sean M. Tweedy, PhD

Abstract: Ordinal-scale strength assessment methods currently used in Paralympic athletics classification prevent the development of evidence-based classification systems. This study evaluated a battery of 7, ratio-scale, isometric tests with the aim of facilitating the development of evidence-based methods of classification. This study aimed to report sex-specific normal performance ranges, evaluate test–retest reliability, and evaluate the relationship between the measures and body mass.

Body mass and strength measures were obtained from 118 participants—63 males and 55 females—ages 23.2 years ± 3.7 (mean ± SD). Seventeen participants completed the battery twice to evaluate test–retest reliability. The body mass–strength relationship was evaluated using Pearson correlations and allometric exponents.

Conventional patterns of force production were observed. Reliability was acceptable (mean intraclass correlation = 0.85). Eight measures had moderate significant correlations with body size ($r =$ 0.30–61). Allometric exponents were higher in males than in females (mean 0.99 vs 0.30).

Results indicate that this comprehensive and parsimonious battery is an important methodological advance because it has psychometric properties critical for the development of evidence-based classification. Measures were interrelated with body size, indicating further research is required to determine whether raw measures require normalization in order to be validly applied in classification.

(*Medicine* 93(4):e31)

Abbreviations: ICC = intraclass correlations, IPC = International Paralympic Committee, MMT = manual muscle testing, SEM = standard error of the mean.

INTRODUCTION

A total of 2.78 million tickets were sold to the London Paralympic Games, making the Paralympic Games the world's third biggest sporting event, after the Olympic games and the FIFA World Cup. The movement is genuinely global —174 countries have National Paralympic Committees—and participation is increasing, with >6000 internationally registered athletes in the sport of athletics alone.[1]

Classification systems are integral to Paralympic sport, being used to determine eligibility and control for the impact of eligible impairment types on the outcome of competition.[2] Valid classification systems facilitate competition in which the athletes who succeed are not simply those who have less severe impairments than their competitors, but those who have the most favorable combination of athletic attributes and have enhanced them to best effect.

Classification that is not valid or that is not perceived to be valid poses a significant threat to Paralympic sport. At the elite level, the legitimacy of an individual's competitive success or athletic achievement can be significantly diminished by the perception that they are in the wrong class, with the potential for considerable personal and financial costs, as well as for discrediting the movement. At the grass-roots level, a classification system that is perceived to be unfair will discourage participation among people with disabilities, rather than achieve the goal of increasing it.

Evidence-based decision making in classification is an essential means of enhancing classification validity, but evidence underpinning current methods of classification is weak. In 2007 the International Paralympic Committee (IPC) mandated the development of evidence-based methods of classification for all Paralympic sports[3] and the IPC Position Stand on Classification outlines key requirements for the development of evidence-based systems, including valid, reliable methods for assessing impairment.[2] Research in this area is limited; however, researchers are beginning to investigate the methods that may contribute to evidence-based classification in Paralympic sport.[4,5]

Impairments classified in Paralympic sport include impaired muscle strength, impaired range of movement, and limb deficiency. The focus of this article is impaired muscle strength, which is required in 16 of the 23 summer Paralympic sports, and is a key component of classification in Athletics (ie, track and field), the largest Paralympic sport. Methods for assessing strength have remained essentially unchanged since the first classification system was described by Sir Ludwig Guttman.[6] They are based on manual muscle testing (MMT) methods in which the strength of individual muscle actions (eg, elbow flexion, knee extension) are assigned a grade from 0 (no voluntary muscle contraction) to 5 (normal strength through normal anatomical range of movement) according to their capacity to overcome gravity and/or manual resistance.[7]

Editor: Shirley S. M. Fong.
Received: April 19, 2014; revised: May 7, 2014; accepted: May 8, 2014.

From the From the School of Human Movement Studies (EMB, MJC, SMT); School of Psychology (PN), University of Queensland, St Lucia, Australia; Department of Rehabilitation Sciences, Faculty of Kinesiology and Rehabilitation Sciences, Catholic University Leuven, Leuven, Belgium (YV).
Correspondence: Emma M. Beckman, University of Queensland, School of Human Movement Studies, Rm 527 Connell Building, St Lucia, Queensland 4072, Australia (e-mail: ebeckman@hms.uq.edu.au).

University of Queensland and Katholieke Universitiet Leuven are IPC Classification Research Partners. This research was supported by the Australian Research Council (grant LP0882187), the International Paralympic Committee, the Australian Sports Commission, and the Australian Paralympic Committee.
The authors have no funding and conflicts of interest to disclose.

ISSN: 0025-7974
DOI: 10.1097/MD.0000000000000031

For the purposes of classification, MMT has a number of advantages: it is widely understood and utilized in clinical practice and, because it does not require any instrumentation, it is inexpensive and space-efficient. MMT also provides a body-size–independent assessment of strength. One final advantage of using MMT for classification is that the ability to overcome manual resistance is assessed using the break test,[8] an isometric contraction against manual resistance. Isometric assessments are advantageous for classification purposes because they are known to be relatively unresponsive to the high-speed, dynamic strength training regimes required for performance enhancement in athletics.[9,10] As a consequence, athletes who have optimized sports performance through high-speed, dynamic resistance training are less likely to change their classification strength measures and be placed into a class for athletes with less severe impairments.[2]

Unfortunately MMT also has several important disadvantages that make it unsuitable as a method of strength assessment for classification. First, acceptable interrater reliability is difficult to achieve, a problem exacerbated by the wide range of MMT techniques that are described in published literature and used by classifiers from different countries in their clinical practice.[11,12] Secondly, the relationship between muscle grade and activity limitation is weak. For example, an athlete with full passive range of motion but only 15 degrees of active elbow extension against gravity is likely to experience much more activity limitation in the shot put than an athlete with 100 degrees active range, and yet the correct muscle grade for both actions is 2.[12] This feature reduces the validity of MMT for the purposes of classification.

More fundamentally however, MMT methods are problematic because they use ordinal scales. Ordinal measurement scales are unsuitable for research that aims to develop evidence-based methods of classification, as mandated by the International Paralympic Committee.[3] Specifically, evidence-based methods of classification require quantification of the relative importance of different muscle actions in a given sporting movement, and therefore a ratio-scale measure of strength is necessary. When the relative importance of key muscle actions has been quantified using a ratio-scale measure, it will be possible to validly aggregate strength measures of contributing muscle groups in order to obtain an evidence-based estimate of how much activity limitation different strength impairments will cause, regardless of their distribution and severity.[2]

In order to permit the development of evidence-based methods of classification, our research group developed a battery of novel strength tests. Key features of the battery were that measures were isometric and therefore, according to theory, less training responsive than other dynamic strength tests; instrumented, yielding an outcome measure in Newtons (a ratio scale); comprehensive, assessing all muscle actions of importance in the key disciplines of Paralympic athletics (wheelchair racing, running, throwing, and jumping); and parsimonious, by assessing compound (or multijoint actions) and thereby minimizing the number of tests required and ensuring that individual tests accounted for the greatest possible variance in performance.[2]

The aims of this study were 3-fold: to establish normal performance ranges for each of the novel tests in nondisabled participants, to evaluate the reliability of each of the novel tests, and to assess the strength of association between individual test outcomes and body mass. In relation to the final aim, a sufficiently strong relationship with body size would indicate that, prior to applying these methods in classification, it may be necessary to develop body-size scaling methods that can be validly applied to measures obtained from athletes with neuromusculoskeletal impairments.

METHODS

Participants

Participants were 118 nondisabled, males (N = 63) and females (N = 55), ages 18 to 37 (mean ± SD 23.2 ± 3.7) recruited from the University of Queensland and local sports clubs. All were regularly active in competitive sport or engaged in 3 or more vigorous training sessions per week. The study was approved by the Ethics Committee of the School of Human Movement Studies, University of Queensland (number HMS07/0406) and all participants provided written informed consent prior to participation. Participants completed testing in a single session, with the exception of 17 who returned in a minimum of 2 days and a maximum of 14 days from the initial session, so that test–retest reliability could be analyzed. The sample size calculation was based on the requirement for calculation of allometric scaling exponents and a minimum number of 54 participants were required in each group (males and females) at an effect size of 0.15 with power set at 0.8 and probability at 0.05.

Strength Testing and Body Size

All participants completed a battery of 7 isometric strength tests—grip strength and 6 novel tests—which are presented in Table 1. The battery aimed to be both parsimonious and comprehensive, using the smallest possible number of tests to evaluate the strength of those movements considered most important in the key athletic disciplines (ie, running, jumping, throwing, and wheelchair propulsion). The order in which participants completed the tests was fully randomized.

Grip strength was assessed using a handheld dynamometer (Smedlay's Dynamometer, Fabrication Enterprises, White Plains, NY, USA, 100 kg), with values entered directly to Excel 2007 spreadsheet (Microsoft, Redmond, WA, USA) and then converted to Newtons (N). The 6 novel tests were completed with the participant seated in a customized strength rig (Figure 1). The rig comprised a rigid, aluminum rectangular frame with an S-type load cell (Scale Components, Slacks Creek, QLD, Australia) rated to 394 kg (1000 lb) mounted at one end, opposite a rigid seat. An aluminium plate (250 mm × 196 mm × 12 mm) was secured to the load cell and, once seated, participants applied force to the load cell from a seated position by either pushing or pulling on the plate. Three features of the rig permitted positioning of participants so that force was applied to the load cell from anatomically standardized positions: the load cell was adjustable vertically and horizontally to account for individual differences in sitting height and breadth (see Figure 1, Panel B); the seat position was adjustable in the fore-aft direction to account for individual differences in arm and leg length (see Figure 1, Panel A); backrest height was adjustable to permit positioning at the C7 vertebra regardless of the participant's sitting height.

The joint angles selected for each testing position aimed to position prime movers so that, as far as possible, length/tension relationships were optimized.[13] To achieve the precise leg angles

TABLE 1. Strength and Body Size Test Protocols

Position	Test	Test Description
Standing	Body mass	Participant stands with weight evenly distributed on scale and mass is recorded to 1 N
Rigid backrest to C7 vertebra; nonelastic strapping securing chest to backrest and pelvis and thighs to seat	Grip strength	Seated position, grip size adjusted for dominant hand, which is held against the trunk with elbow flexed at 90 degrees. Performed with dominant and nondominant hands
	Single Supported arm push	Hand of testing arm positioned on push plate, wrist raised to shoulder level and aligned in sagittal plane (shoulder at 90 degrees abduction, 45 degrees horizontal shoulder flexion, 120 degrees elbow extension). Nontesting arm resting in lap. Performed with dominant and nondominant arm
	Bilateral supported arm push	Both hands on push plate with wrists raised to shoulder level, aligned sagittally with centre of sternum (90 degrees abduction, 45 degrees horizontal shoulder flexion, 120 degrees elbow extension). This test is illustrated in Figure 2A
	Unsupported push/Pull	Sitting on seat, independent of backrest. Hand of dominant arm on push plate with wrist raised to shoulder level and aligned sagittally (shoulder at 90 degrees abduction, 45 degrees horizontal shoulder flexion, 120 degrees elbow extension). Nondominant hand grips rigid pole at shoulder height
Backrest to C7 vertebra; nonelastic strapping securing trunk/pelvis to backrest	Leg flexor strength	Foot of testing leg on foot on push plate (1200 degrees knee extension, 700 degrees hip flexion). Foot of nontesting leg on ground
	Leg extensor strength	Foot of testing leg secured with nonelastic strapping to push plate (120 degrees knee extension, 600 degrees hip flexion). Foot of nontesting leg on ground. This test is illustrated in Figure 2B
	Plantar flexor strength	Foot of testing leg on push plate (Full knee extension, ankle starts in neutral to 5^0 dorsiflexion). Foot of nontesting leg on ground.

described in Table 1, an anthropometric marking pen was used to clearly identify the following landmarks: acromion, greater trochanter, mid-point of lateral knee joint line, and the lateral malleolus. A Sony camera, positioned 1.5 m from the rig, in line with the mid-point of the seat surface at 1 m high, fed a live, sagittal view video image of the participant seated in the strength rig into a personal computer, on which Dartfish (version 4.0.9.0; Dartfish, Lausanne, Switzerland) was installed. The Dartfish angle tool was used to draw the required joint angles and the strength rig could then be adjusted until the participant was positioned so that their video image was matched with image created by the angle tool. The live feed permitted monitoring to ensure required positioning was maintained. This method could not be used for the arm angles which are only

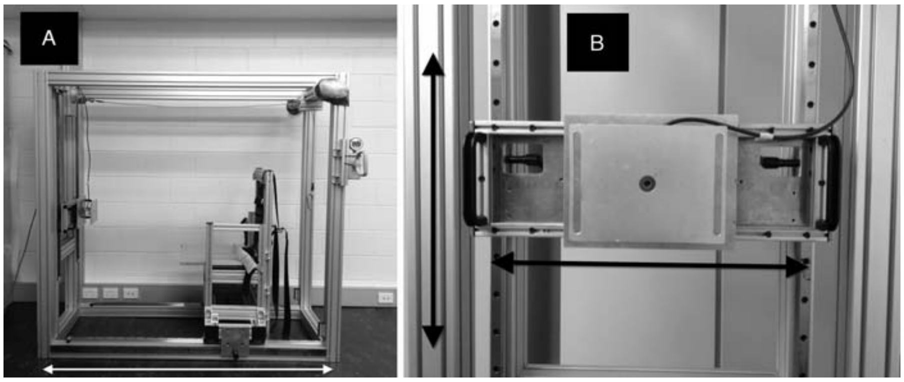

FIGURE 1. Panel A shows the rigid aluminum frame in which the strength tests were conducted. The arrow indicates the ability of the chair to move in the fore-aft direction. Panel B shows a magnified view of the load cell set up with arrows indicating the ability to move the load cell vertically and horizontally.

FIGURE 2. Panel A illustrates the position that participants were placed into for the bilateral Supported arm push. Panel B illustrates the position that participants were placed into for the leg extensor strength test.

observable from an overhead view, and these angles were set using a handheld goniometer (SunShine Diagnostic and Measuring Instruments, New Taipei City, Taiwan) and monitored visually by a member of the testing team. Figure 2 shows athletes positioned for an upper body strength measure (Panel A) and a lower body strength measure (Panel B).

Once positioned, participants performed 3 maximal isometric contractions of 5 seconds duration, each separated by 30 seconds of rest. Valid trials required peak force to be achieved slowly (>2 seconds and <3 seconds) followed by a 3-second hold at peak. To assist participants achieve peak force slowly, 2 submaximal practice trials were performed with real-time visual feedback of the force–time curve so that participants could easily identify when maximum force was achieved either too quickly or too slowly.

For all tests participants were instructed to "push the plate straight back into the wall." Participants applied force with their hand(s) for upper limb tests and with their foot/feet for lower limb tests, with the rigid backrest permitting exertion of maximal voluntary force. Hands/feet were positioned so that force was directed through the long axis of the load cell and torque was minimized.

Load cell output was captured by Musclelab v4020e (Ergotest, Porsgunn, Norway) at a sampling rate of 100 Hz. Raw isometric strength scores were exported directly into an Excel 2007 spreadsheet from the load cell data acquisition program (Musclelab), and all trials were then processed in SciLab using a custom-written algorithm (Scilab Enterprises, Versailles, France) to acquire the peak isometric force from each trial. For each trial, the isometric force used for further analysis was calculated as the mean force over the 2-second period with the least variability. To ensure a plateau was achieved, a trial was deemed acceptable if the calculated force was a minimum of 95% of the peak force registered during the trial. The best trial was then employed for statistical analysis.

Body size was assessed through the measurement of body mass. Body mass was measured and recorded on each participant to the nearest 0.1 kg using scales (Seca 760 Mechanical Scales, Seca, Hamburg, Germany).

Data Analysis

All data were analyzed using SPSS v16 (SPSS Inc, Chicago, IL). The data were assessed for normality using the Shapiro–Wilk test of normality. Independent *t*-tests were performed on the strength data of males and females to determine whether strength tests were significantly affected by sex and should be analyzed separately. Test–retest reliability was assessed using dependant *t*-tests, 2-tailed intraclass correlations (ICC) method 3,1, standard error of the mean (SEM) and Bland–Altman plots.

The relationship between body mass and each of the strength measures was determined by calculating Pearson correlations and allometric scaling exponents. To calculate the latter, a log-linear regression analysis was conducted using strength as the dependent variable and body mass as the independent variable. The slope of the regression line was used as the allometric scaling exponent,[14] and residual plots were assessed to check the fit of each model. Exponents were calculated on the dominant side for the upper body and the right side for the lower body.

RESULTS

Isometric strength test data were normally distributed. Independent *t*-tests showed significant differences between males and females in all strength measures, and therefore

TABLE 2. Descriptive Statistics for Body Mass and Strength Measure

		Males				Females			
Test Name	**Limb**	**n**	**Mean (N)**	**SD**	**Range**	**n**	**Mean (N)**	**SD**	**Range**
Body mass	N/A	62	756.3	94.8	605–985	54	634.3	82.6	510–800
Grip strength	Dominant	62	515.8	8.52	38.0–73.0	55	344	5.2	24.0–47.0
	Nondominant	62	495.5	8.75	29.0–72.0	55	323.6	4.5	23.0–43.0
Single supported arm push	Dominant	61	465.0	134.2	192.4– 823.5	54	271.3	69.5	134.6–414.9
	Nondominant	62	473.1	129.4	211.5–772.4	55	292.3	87.8	173.9–598.3
Bilateral supported arm push	N/A	62	1017.4	292.7	424.2–1661.2	54	547.6	144.5	318.9–899.3
Unsupported push pull	N/A	58	386.2	112.4	154.5–629.0	48	224.0	57.0	98.0–326.4
Leg flexor strength	Left	58	339.2	79.9	167.8–487.7	50	225.8	59.3	93.1–375.3
	Right	58	352.9	78.3	124.0–556.5	51	237.4	53.6	93.1–352.7
Leg extensor strength	Left	59	1786.1	486.5	883.0–2823.1	52	1136.1	299.2	464.5–1594.5
	Right	58	1822.4	454.6	884.0–2701.0	52	1193.5	302.1	568.5–1768.7
Plantar flexor strength	Left	50	1464.2	364.1	651.2–2201.8	51	1128.5	326.7	510.1–1990.7
	Right	51	1443.2	345.6	814.5–2321.6	51	1124.3	337.8	588.1–2147.2

these data were analyzed separately. Table 2 presents descriptive statistics for males and females for isometric strength measures. Overall, both male and female participants were symmetrical, with no significant difference in mean force production between dominant and nondominant arms, left and right leg flexors, and left and right leg extensors. In the lower limbs, extensor strength (male mean right leg: 1822.4N ± 454.6N; female mean right leg: 1193.5N ± 302.1N) was approximately 5 times greater than flexor strength (male mean right leg: 352.9N ± 78.3N, female mean right leg: 237.4N ± 53.6N).

The results of analyses for test–retest reliability are presented in Table 3. There was no systematic bias in the test–retest results as indicated by a nonsignificant *t*-test. For all tests, ICCs ranged from 0.71 to 0.95. The mean difference between the test and retest was consistently low for all tests—absolute difference range was 3.7 to 51.0 N, a relative difference of 0.2% to 7.3%. SEM values were also low and considered acceptable for all tests (18.12–117.44).

Pearson correlations showing the relationships between individual strength tests and body mass are presented in Table 4 for males, females, and the total sample. Correlations were moderate and significant for 6 of the 7 tests in males ($r = 0.43$–0.61, $P = 0.000$) and 2 of the 7 tests in females ($r = 0.30$–0.33, $P = 0.014$–0.03).

Allometric exponents were calculated separately for males and females; exponents were higher for males (range = 0.58–1.27) than for females (range = 0.06–0.67). Residual plots showed all models were appropriately fitted.

DISCUSSION

Currently, the method of strength assessment used in Paralympic athletics classification is based on MMT, an ordinal-scale measure. This is problematic because ordinal-scale measures do not permit quantitative evaluation of the impact of impairment on athletic performance, a fundamental requirement for the development of evidence-based methods of classification. The results from this study indicate that the ratio-scale strength assessment battery described could be validly applied to address this important methodological shortcoming and facilitate the development of evidence-based methods of classification. In addition to furnishing a ratio-scale measure of strength, the battery has a number of features that would be advantageous for Paralympic classification, including the measures are isometric and therefore training resistant; the battery is both comprehensive (assessing resultant forces produced by the muscle groups of principal importance in Paralympic Athletics) and parsimonious; and reliability is good to excellent. Furthermore, the

TABLE 3. Test–Retest Reliability of Novel Strength Measures (n = 17)

Test Name	T1 M(SD) (N)	T2 M (SD)* (N)	ICC (95% CI)	SEM (N)	Mean Δ T1 – T2 (N)
Single Supported arm push	221.3 (59.5)	229.9 (70.1)	0.93 (0.6–0.95)	30.0	8.6
Bilateral supported arm push	466.6 (109.8)	432.8 (104.5)	0.81 (0.44–0.93)	60.9	–33.9
Unsupported push pull	196.3 (64.0)	191.1 (52.1)	0.95 (0.85–0.98)	18.1	–5.2
Leg flexor strength	230.9 (43.2)	224.3 (50.0)	0.71 (0.17–0.90)	31.4	–6.6
Leg extensor strength	1009.1 (219.4)	1005.4 (200.9)	0.80 (0.41–0.93)	117.4	–3.7
Plantar flexor strength	989.0 (241.1)	964.1 (294.5)	0.92 (0.77–0.97)	103.7	–24.9

ICC = intraclass correlations, SEM = standard error of the mean, SD = standard deviation.
*No significant difference was found between the mean force production in T1 and T2.

TABLE 4. Pearson Correlations and Allometric Scaling for Body Mass and Strength Measures

	Males Pearson Correlation	Allometric Exponent	Females Pearson Correlation	Allometric Exponent
Grip strength	0.46**	0.58	0.33*	0.40
Single supported arm push	0.51**	1.24	0.11	0.31
Bilateral supported arm push	0.45**	1.04	0.15	0.35
Unsupported push/pull	0.43**	1.03	0.11	0.67
Leg flexor strength	0.06	0.62	0.17	0.06
Leg extensor strength	0.59**	1.27	0.30*	0.21
Plantar flexor strength	0.61**	1.19	0.22	0.10

*Correlation is significant at the 0.05 level.
**Correlation is significant at the 0.01 level.

sex-specific normal performance ranges that are reported will permit meaningful interpretation of results in athletes with impairments in future studies. The key advantages of the battery evaluated in this study are expanded in the following paragraphs.

The following example illustrates why ratio-scale measurement is of fundamental importance to the development of evidence-based methods of classification. In wheelchair racing, the current class T51 is based on the strength impairment profile of a person with a complete spinal cord injury at neurological level C5–6, including impaired shoulder flexion strength (up to grade 4) and triceps strength grade 0–3.[7] The current class T52 profile is based on the strength impairment profile of a person with complete spinal cord injury at neurological level C7–8, including normal shoulder flexion and triceps strength (ie, grade 5).[7] If an athlete with polio presents with normal shoulder flexion strength but grade 3 triceps, the classifier must decide whether the athlete is best in class T51 (even though they will have more shoulder flexion strength than other athletes in that class) or in class T52 (even though they will have less triceps strength than other athletes in class). In this instance, decision-making should be based on evidence regarding the relative importance of shoulder flexion and triceps extension to wheelchair propulsion. Current, ordinal-scale methods of strength assessment are not suitable for investigating this question.

Two additional advantages conferred by the proposed test battery are that it is both parsimonious and comprehensive. These advantages occur because, rather than assessing individual muscle actions acting over a single joint, a number of the proposed tests assess the resultant force of a number of key muscle groups acting over more than one joint. For example, the leg extension test assesses the combined strength of hip and knee extensors simultaneously, while the single Supported arm push assesses shoulder horizontal flexion and elbow extension simultaneously. Consequently, the test battery described in this study comprises 7 tests, rather than the 20 individual muscle grade tests required to assess the same muscle groups in the current system. In addition to saving time, the proposed test battery would considerably reduce the number of maximum contractions required from each athlete during classification, helping to ensure that fatigue did not the confound outcomes.

Importantly, the significantly reduced number of tests does not result in a less comprehensive battery—all the principal muscle actions required for the activities that are central in Paralympic athletics are evaluated (ie, wheelchair racing, running, jumping, and throws—both standing and seated). More specifically, the tests capture muscle synergies that are required for performance of the activities of interest. For example, the push pull test for the upper body uses the same prime movers as used for seated throwing with a pole: shoulder extension and elbow flexion of the nonthrowing arm simultaneously assessed at the same time as shoulder flexion and elbow extension on the throwing or dominant arm.

The sex-specific isometric strength ranges reported in this paper on nondisabled individuals will allow the interpretation of results obtained from athletes with impairments. In general, the overall pattern of results was consistent with what is known about strength—males were significantly stronger than females on all strength tests; the lower body was stronger than the upper body and the lower limb extensor strength was greater than the flexor strength. More specifically, Grip strength means (males = 526 N; females = 344 N) were comparable with means previously reported for this protocol (males = 523 N; females = 319 N)[15] and the mean bilateral supported arm push for males (N) was similar to that reported by Hortobagyi (9976 N) for a similar protocol performed in supine.[16] However, the battery also extends what is known about the relative strength of different movements—for example, combined hip and knee flexion was only 19.4% of extension in males and 19.9% in females. These percentages are considerably less than those reported in studies of isolated lower limb movements, which indicate that knee flexion strength is between 43% and 90% of knee extension strength.[17,18]

One final feature of the test battery described in this study which is important for the purposes of classification is that the reliability of all strength tests was excellent (ICC > 0.8) for all but one test.[19] The ICC for the leg flexor test was good with an ICC of 0.71,[19] which is acceptable given the low SEM (314 N) and the small mean difference between test one and test 2 (66 N).

As mentioned in the introduction, one advantage that current methods of strength assessment have is that they are independent of body size. This is an important feature because classification aims to control for the impact of strength impairment on athletic performance without controlling for other advantages conferred by body size. For example, 2 throwers with complete spinal cord injuries at T2 should compete in the same class for the discus, regardless of whether one is 2 m tall and weighs 100 kg and the other is

1.5 m tall and 65 kg. It will be important that any new measure of strength used for the purposes of classification is, as far as possible, independent of body size.

Results from this study indicate that a number of measures—1 in males (leg flexor strength) and 5 measures in females (single supported arm push, bilateral supported arm push, unsupported push pull, leg flexor and plantar flexor strength)—were not strongly or significantly related to body mass. This indicates that it is likely that these measures are sufficiently independent of body size to be validly applied in classification. However, 6 tests in males were moderately and significantly related to body mass ($r = 0.43$–0.61) and 2 tests in females showed a weaker but still significant relationship (0.30–0.33). Allometric exponents were also calculated to give an indication of the slope of the regression line as this would indicate how much change in strength would be expected for a given change in body mass. The allometric exponents that were calculated were much larger for males (mean of 1.06) than for females (mean of 0.3). This indicates that a given change in body size is associated with a larger change in strength in males than in females. It has been suggested by previous research that this may be due to a decrease in range of strength scores seen when males and females are analyzed separately,[16,20] or a result of higher relative percentages of lean body mass in males when compared with females.[15,21] These results indicate that further research is required to determine whether raw measures require normalization in order to be validly applied in classification.

Unfortunately the scaling exponents developed in this study on nondisabled participants based on body mass will not be able to be applied directly to athletes with disabilities. Body measurements such as body mass, height, or limb circumference are inappropriate in athletes with disabilities as neuromusculoskeletal impairment changes the relationship between these measures of body size and strength in ways that are unpredictable. This is because impairments to the central and peripheral nervous system and to the muscle fiber itself, which commonly affect Paralympic athletes, will disrupt the fundamental premise for scaling—that force is primarily determined by muscle cross-sectional area. For example, people affected by spastic paraplegia resulting from UMN injury will retain muscle bulk better than those affected by lower motor neuron injuries due to the presence of intact spinal-level reflexes, although the impairment of structures may result in comparable impairments of strength function.

It is posited that local bony dimensions—for example, humerus length or biacromial width—may be the most appropriate type of anthropometric measure by which to scale strength measures in athletes with neuromusculoskeletal impairment because, compared to body mass or muscle cross-sectional area, they more commonly remain unaffected by neuromusculoskeletal impairments. Research is required that identifies the most appropriate local bony dimensions by which to scale [22] and subsequently evaluates whether normalized or raw scores are more valid for classification purposes.

CONCLUSIONS

Ordinal-scale strength assessment methods are currently used in Paralympic athletics classification, preventing the development of evidence-based classification systems. The results from this study indicate that the ratio-scale strength assessment battery described could be validly applied to address this important methodological shortcoming and facilitate the development of evidence-based methods of classification by allowing a valid and reliable assessment of muscle strength in athletes with neuromusculoskeletal impairments. The battery has a number of other features that are advantageous for classification, including the measures are isometric and therefore training resistant; the battery is both comprehensive (assessing resultant forces produced by the key muscle groups) and parsimonious; and reliability is good to excellent. Furthermore, the sex-specific normal performance ranges that are reported will permit meaningful interpretation of results in athletes with impairments in future studies. The results from this study have implications for the 16 Paralympic sports that assess strength in the classification process.

REFERENCES

1. *Annual Report*. Bonn, Germany: International Paralympic Committee, 2012.
2. Tweedy SM, Vanlandewijck YC. International Paralympic Committee Position Stand: background and scientific rationale for classification in paralympic sport. *Br J Sports Med*. 2011;45:259–269.
3. *IPC Classification Code September 2007*. Bonn, Germany: International Paralympic Committee, 2007:15.
4. Beckman EM, Tweedy SM. Evaluating the validity of activity limitation tests for use in paralympic classification. *Br J Sports Med*. 2009;43:1067–1072.
5. Oh Y, Burkett B, Osboirough C, et al. London 2012 Paralympic Swimming: passive drag and the classifciation system. *Br J Sports Med*. 2013;47:838–843.
6. Guttman L. *Textbook of Sport for the Disabled*. St Lucia, Australia: University of Queensland Press, 1977.
7. *IPC Athletics Classification Rules and Regulations*. Bonn, Germany: International Paralympic Committee, 2013.
8. Daniels L, Worthingham C. *Muscle Testing Technique of Manual Examination*. 5th ed. Philadelphia: WB Saunders Co, 1986.
9. Cronin J, Hansen KT. Strength and power predictors of sports speed. *J Strength Cond Res*. 2005;19:349–357.
10. Wilson GJ, Lyttle AD, Ostrowski KJ, Murphy AJ. Assessing dynamic performance: a comparison of rate of force development tests. *J Strength Cond Res*. 1995;9:176–181.
11. Bohannon R. Manual muscle testing does it meet the standards of an adequate screening test? *Clin Rehabil*. 2005;19:662–667.
12. Tweedy SM, Williams G, Bourke, J. Selecting and modifying methods of manual muscle testing for classification in Paralympic sport. *Eur J Adapted Physical Activity*. 2010;3:7–16.
13. Murphy A, Wilson GJ, Pryor JF, Newton RU. Isometric assessment of muscular function: the effect of joint angle. *J Appl Biomech*. 1995;11:205–215.
14. Jaric S. Muscle strength testing. Use of normalisation for body size. *Sports Med*. 2002;32:615–631.
15. Vanderburgh PM, Mahar MT, Chou CH, et al. Allometric scaling of grip strength by body mass in college-age men and women. *Res Q Exerc Sport*. 1995;66:80–84.
16. Katch F, LaChance P, Behnke A, et al. Relationships of body size, segmental dimensions, and ponderal equivalents to muscular strength in high-strength and low-strength subjects. *Int J Sports Med*. 1990;11:349–356.
17. Markhede G, Grimby G. Measurement of strength of hip joint muscles. *Scand J Rehabil Med*. 1980;12:169–174.

18. Nosse L. Assessment of selected reports on the strength relationship of the knee musculature. *J Orthop Sports Phys Ther*. 1982;4:78–85.
19. Fleiss JL. *The Design and Analysis of Clinical Experiments*. New York: John Wiley & Sons, 1986.
20. van den Tillaar R, Ettema G. Effect of body size and gender in overarm throwing performance. *Eur J Appl Physiol*. 2004;91:413–418.
21. Bazett-Jones D, Cobb SC, Joshi MN, et al. Normalizing hip muscle strength: establishing body-size-independant measurements. *Arch Phys Med Rehabil*. 2011;92:76–82.
22. Maeda T, Oowatashi A, Kiyama R, et al. Prediction of muscle strength using length and width of bone. *J Phys Ther Sc*. 2001;13:27–30.