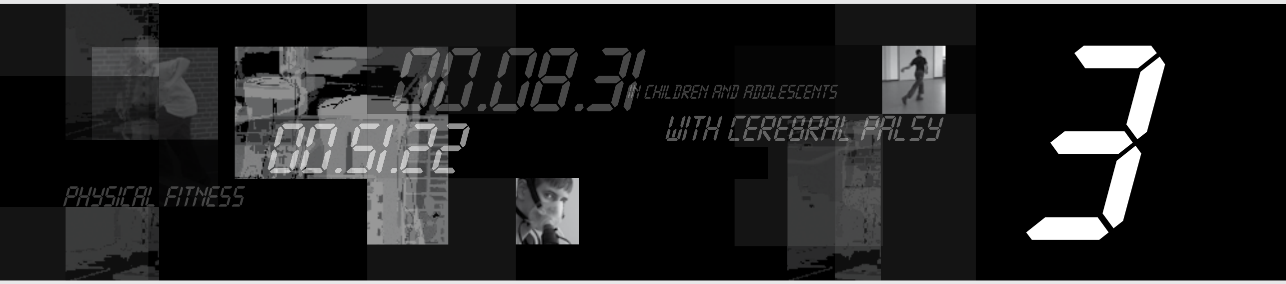

RELIABILITY AND VALIDITY OF DATA FOR 2 NEWLY DEVELOPED SHUTTLE RUN TESTS IN CHILDREN WITH CEREBRAL PALSY

Olaf Verschuren, Tim Takken, Marjolijn Ketelaar,
Jan Willem Gorter, Paul J.M. Helders

Physcial Therapy 2006;86:1107-1117

Abstract

Background and Purpose The purpose of this study was to examine the reliability and validity of data obtained with 2 newly developed shuttle run tests (SRT-I and SRT-II) to measure aerobic power in children with cerebral palsy (CP) who are classified at level I or II on the Gross Motor Function Classification System (GMFCS). The SRT-I was developed for children at GMFCS level I, and the SRT-II was developed for children at GMFCS level II.

Subjects Twenty-five children and adolescents with CP (10 female, 15 male; mean age=11.9 years, SD=2.9), classified at GMFCS level I (n=14) or level II (n=11), participated in the study.

Methods To assess test-retest reliability of data for the 10-m shuttle run tests, the subjects performed the same test within 2 weeks. To examine validity, the shuttle run tests were compared with a GMFCS-level based treadmill test designed to measure peak oxygen uptake.

Results Statistical analyses revealed test-retest reliability for exercise time (number of levels completed) (intraclass correlation coefficients of .97 for the SRT-I and .99 for the SRT-II) and reliability for peak heart rate attained during the final level (intraclass correlation coefficients of .87 for the SRT-I and .94 for the SRT-II). High correlations were found for the relationship between data for both shuttle run tests and data for the treadmill test (both r=.96).

Discussion and Conclusion The results suggest that both 10-m shuttle run tests yield reliable and valid data. Moreover, the shuttle run tests have advantages over a treadmill test for children with CP who are able to walk and run (GMFCS level I or II).

Introduction

Maximal aerobic power is an important component of physical fitness and often is used to measure the effect of physical exercise training. It also is used to assess cardiovascular capacity as well as the level of fitness in children and adolescents with cerebral palsy (CP).[1] Ascertaining maximal aerobic power is the most common way to assess the outcome of physical exercise training[2] in children with CP. Peak oxygen uptake (VO_{2peak}) determined at the limits of exhaustive treadmill or cycle exercise is generally viewed to be the best physiologic marker of aerobic fitness.[3]

Currently, tests for assessing aerobic power in children with CP that are clinically practical and easy to use and that yield reliable and valid data are absent. Most of the available literature has focused on assessments of VO_{2peak} using cycle ergometry[4] or arm ergometry.[5] Treadmill running engages a larger muscle mass than does cycling[6]; therefore, VO_{2peak} is typically 8% to 10% higher during treadmill running than during cycle ergometry.[7] A major disadvantage of cycle ergometry in children is that a high proportion of the total power output is generated by the quadriceps femoris muscle and the effort required to push the pedals during the later stages of a progressive test is high in relation to muscle strength (force-generating capacity).[8] Children with CP have weak quadriceps femoris muscles.[9] The maximal performance during a cycle ergometry test, therefore, can be limited by a peripheral factor. Still, many researchers reported VO_{2peak} of children with CP using the cycle ergometer.[4,10,11]

For children who are able to walk independently, the most functional way to assess their maximal aerobic power would be a walking- or running-based exercise test. The treadmill protocols that currently are often used in clinical practice (Bruce protocol and Balke protocol) are not appropriate for children with CP.[1] For most children with CP who have problems with movement coordination and an equinus position of the foot, the increasing speed and inclining floor are problematic and difficult. Research, however, has shown that it is possible for children with CP to perform a maximal exercise test on a treadmill,[12,13] although there is no published standardized protocol for administering VO_{2peak} tests to those children.

Although the "gold standard" assessment of exercise tolerance in children can be measured in a laboratory using a treadmill, the necessary equipment is expensive and may not be readily accessible. Thus, substantial logistical problems must be overcome before a child with CP can be assessed on the treadmill, and there is a clinical need to develop and validate field tests of exercise tolerance in children and adolescents with CP.

Shuttle run tests are field tests in which a participant walks or runs between 2 markers. These tests are potentially useful measures of exercise tolerance. Because there is no need for expensive equipment, shuttle run tests can easily be administered in a clinical setting. The shuttle run test commonly used is the

20-m shuttle run, which was developed and validated by Leger et al.[14] This test has been demonstrated to be an effective measure of the aerobic fitness in subjects who were healthy.[15,16] Validity has been determined in children who are developing typically,[14,17,18] in athletes,[19] and in a sedentary adult population.[20]

For most children and adolescents with CP, the 20-m shuttle test is not suitable, because the starting speed (8 km/h) and the increase (0.5 km/h) every minute are beyond their capabilities. Several experiments with the 20-m shuttle test showed that most of the children and adolescents with CP who were able to walk independently were not able to complete the first level of the test or reached exhaustion before 5 minutes. A continuous progressive exercise lasting 8 to 17 minutes is optimal for achieving a maximal myocardial oxygen demand for diagnostic and prognostic purposes.[21]

To facilitate assessment in children, the tests should be nonthreatening, inexpensive, and easy to administer.[22] For this reason, we developed 2 treadmill protocols and shuttle run tests for children with CP who are classified at level I or II on the Gross Motor Function Classification System (GMFCS).[23] To achieve a maximal exercise duration between 8 to 17 minutes, the GMFCS classification seems to be a useful system to distinguish between the 2 groups: children with a level I classification, who are able to run, and children with a level II classification, who encounter more difficulties while running. Therefore, we developed 2 different protocols with different starting speeds.

In this study, we examined the test-retest reliability and validity of data obtained with a GMFCS level I-specific 10-m shuttle run test (SRT-I) and a GMFCS level II-specific 10-m shuttle run test (SRT-II). We also evaluated the feasibility of both tests for use with children with CP.

Table 1. Subject Characteristics (n=25)

	GMFCS[a] Level I (n=14)				**GMFCS Level II (n=11)**			
Variable	**Mean**	**SD**	**Median**	**Range**	**Mean**	**SD**	**Median**	**Range**
Age (y)	11.5	2.8	11.7	7.5-16.1	12.5	3.0	12.1	7.2-17.0
Height (cm)	148.7	15.3	149.0	125.0-175.0	148.6	18.9	145.0	123.0-175.0
Body mass (kg)	40.3	12.4	35.1	23.8-60.8	38.6	12.1	32.7	24.0-59.7
Sum of 7 skinfold measurements (mm)	77.2	44.3	61.8	36.0-196.5	74.2	28.5	70.5	33.0-131.5

[a]GMFCS=Gross Motor Function Classification System.

Materials en Methods

Subjects

A convenience sample of 30 children and adolescents from a school for special education were invited to participate in the study. To be included, subjects were required to be within the age range of 7 to 20 years, had to be diagnosed with CP, and classified at level I or II on the GMFCS.[23] Cognitively, they had to be capable of following simple commands. Twenty-five subjects (15 male, 10 female) and their parents agreed to participate and signed an informed consent form. Group characteristics according to GMFCS level are described in Table 1.

Procedure

Prior to testing, each child was weighed on electronic scales (Seca, Hamburg, Germany). Height measurements were taken on the same visit while the child was standing against a wall. Body coposition was assessed using the sum of 7 skinfold measurements according to the method described by Pollack et al.[24] The skinfold measurements were taken at 7 sites on the right side of the body (triceps, biceps, subscapular, suprailiac, mid-abdominal, medial calf, and front thigh) by 2 of the investigators (OV and TT) in accordance with the American College of Sports Medicine guidelines.[25]

Study Design

To assess test-retest reliability of data for the 10-m shuttle run tests, each subject performed one shuttle run test (SRT-I or SRT-II) 2 times. To assess the validity of data for both shuttle run tests, the VO_{2peak} values obtained with gas analysis during the GMFCS level-based treadmill test and the shuttle run tests were compared. The treadmill test and the shuttle run tests were separated by a minimum of 2 days and a maximum of 7 days ($\overline{X}$=4.1, SD=1.4).

Treadmill testing on all subjects was performed using protocols that we developed. Subjects with a level I or II classification on the GMFCS performed different protocols. All subjects practiced walking on the treadmill before testing for a maximum of 3 minutes. The SRT-I was used for the subjects who were classified at level I on the GMFCS, and the SRT-II was used for the subjects who were classified at level II on the GMFCS.

To assess validity, we compared the VO_{2peak} achieved on the treadmill and the VO_{2peak} achieved on the 10-m shuttle run test. During the study, all subjects performed 1 treadmill test and 2 identical shuttle run tests within 2 weeks to assess the reproducibility of the measurements. One shuttle run test and the treadmill test were performed while the subjects were wearing a facemask for gas analysis. The subjects also completed one shuttle run test without a facemask and consequently no gas analysis. One child did not wear a facemask because the thought of breathing through the mouthpiece frightened her, and only the heart rate was monitored during all tests for that subject. The testing was done while the subjects were wearing regular sportswear and shoes, and orthoses if applicable.

In the tests followed by gas analysis, the subjects wore a firmly fitted facemask attached to a calibrated mobile gas analysis system (Cortex Metamax B3)[†] with an in-built gas analyzer, which allowed breath-by-breath gas analysis throughout the tests. Measurements of cardiopulmonary variables were collected. Before each trial started, the subjects rested until 1 minute had passed without an increase in heart rate. The subjects' heart rate was measured continuously, and the measurements were saved to a storage device during all tests using a reliable and accurate heart rate monitor.[‡26] The Cortex Metamax is a valid and reliable system for measuring ventilatory parameters during exercise.[27-29]

The mobile gas analysis system consisted of a facemask, a transmitting unit (containing different oxygen and carbon dioxide gas analyzers), and a receiving unit. The transmitting unit with facemask with tubing (total weight=0.57 kg) was attached to the subjects with a harness, and the receiving unit was connected to a laptop computer located within 500 m of the transmitting unit. Metabolic stress test software (Metasoft, Version 2.6) was used to measure breath-by-breath minute ventilation, oxygen consumption (VO_2), carbon dioxide production (VCO_2), and heart rate and to calculate the respiratory exchange ratio ($RER=VCO_2/VO_2$). During all tests, the subjects were verbally encouraged to run as long as possible.

During the shuttle run tests without gas analysis, only the subjects' heart rate was monitored. The heart rate was read from the wrist monitor at the end of the test and recorded on the datasheet.

Measures and Treadmill Test and 10-m Shuttle Run Test Protocols

GMFCS. The GMFCS, translated into the Dutch language, was used by a pediatric physical therapist (OV), who was experienced in using the GMFCS, to classify the children and adolescents with CP into groups based on their functional ability. Level I represents the highest level of functional abilities, and level V represents the lowest level of functional abilities. Due to the physical demands of the tests, only children and adolescents who were classified at GMFCS level I (able to walk indoors and outdoors and climb stairs without limitation) or level II (able to walk indoors and outdoors and climb stairs holding on to a railing, but experience limitations in walking on uneven surfaces and inclines and in walking in crowds or confined spaces) were recruited. The original GMFCS has been reported to yield reliable and valid data for children aged 6 to 12 years.[23] Children over 12 years of age were classified using the same criteria as those used for 6- to 12-year-olds.

[†] Cortex Medical GmBh, Leipzig, Germany

[‡] Polar, Kempele, Finland

Treadmill tests. We could not adhere to "classic" treadmill protocols (Bruce protocol or Balke protocol) because they are not appropriate for children with CP due to the gait disturbance and accompanying problems. We performed a pilot study to develop the new test protocols. The starting speed and increase in speed were adjusted until mean total exercise time was between 8 and 17 minutes.[21] After this pilot study, we developed 2 treadmill protocols: 1 protocol for children who are classified at level I on the GMFCS and 1 protocol for children who are classified at level II on the GMFCS.

Our protocols started at 5 km/h for the GMFCS level I treadmill test and at 2 km/h for the GMFCS level II treadmill test. Speed was increased 0.25 km/h every minute by the assessor. Jones and Doust[30] showed that at the 2 lowest speeds, VO_2 during road running was not significantly different from treadmill running at low grades. Therefore, throughout the test we used a slope of 2% to compensate for the lack of air resistance, which results in a lower energy cost. With this incline on the treadmill, the results can be compared with those obtained during indoor running at the same speed.[30]

When walking, the subjects were permitted to support themselves with their fingers on the guardrails of the electromechanical treadmill (ENMil)[§] to maintain their balance. Subjects with a level II classification on the GMFCS performed their exercise tests on a wheelchair treadmill.[§] This treadmill is modified to accommodate the width of a wheelchair and is equipped with side guards. Children with a broad-based gait and other gait disturbances can stay more easily on this kind of treadmill. To ensure that the subjects were safe during the test, a therapist assisted them while walking behind or next to them. An emergency button was always within reach of the subjects and the therapist.

A maximal exercise test on the treadmill was done after a 3-minute treadmill practice session to familiarize the subjects with the equipment. The speed during this session was 2 km/h, and the practice session was followed by a 5-minute rest. The subjects were encouraged to push themselves to their limits, and the test was stopped when the subjects were unable or refused to continue the test despite encouragement.

All tests were performed under standardized conditions in a laboratory environment. The subjects maintained their normal diet before the day of testing. Only light physical activity was performed on the day before testing, and on the test day, subjects did not exercise before their test visit. During their visit, the subjects were given adequate explanation of the proposed protocol and its objectives.

[§] Eraf, Delft, The Netherlands

10-m shuttle run tests. Two new shuttle run tests for children with a GMFCS level I or II classification were developed. The starting speed and the increase in speed every minute are the same as for the treadmill protocol—5 km/h for the SRT-I and 2 km/h for the SRT-II—with the speed increased 0.25 km/h every minute.

Both shuttle run tests require children to walk or run between 2 markers delineating the respective course of 10 m, at a set incremental speed determined by a signal, which is played by a standard CD player. All subjects were accompanied by a physical therapist during the test to help them pace themselves with the audio signal. At the end of each level, the subjects were told to go a little faster. The test was finished when, on 2 consecutive paced signals, the subjects were more than 1.5 m away from the marker. Total exercise time was recorded and used for analysis.

The treadmill test and the shuttle run test with gas analysis were supervised by 2 experienced assessors (OV and TT). The shuttle run test without gas analysis was supervised by assessor (OV) and one pediatric physical therapist who was randomly chosen out of a sample of 6 therapists with no experience or formal training. One therapist was instructed to encourage the child, and 1 therapist (OV) accompanied the child during the test.

Feasibility

At the completion of the tests, subjects were asked a standardized question: "Which of the 2 tests, shuttle run or treadmill, did you prefer, and why?" All of the answers were recorded by the investigator for analysis.

Data Analysis

The data were analyzed using SPSS 12.0 and MS Excel 2003 for Windows. Intraclass correlation coefficients (ICC[2-way mixed]) for the number of levels completed were computed to assess test-retest reliability of data for both 10-m shuttle run tests. Acceptable reliability was considered to be an ICC value greater than .80.[31]

Limits of agreement also were calculated according to the procedure described by Bland and Altman.[32] A Bland-Altman plot is a graphic representation of the individual subject differences between the tests plotted against the respective individual means. Using this plot rather than the conventional test-retest scattergram, a rough indication of systematic bias and random error is provided by examining the direction and magnitude of the scatter around the zero line, respectively. Bland-Altman analysis describes the level of agreement between 2 measurements. In this analysis, the "precision" indicates how well the methods agree for an individual. By multiplying the precision by 1.96, the "limits of agreement" are calculated. This calculation represents the 95% likely range for the difference between a subject's scores on 2 tests and is an indicator of absolute reliability. Typical error and total error were calculated as described by Hopkins.[33] Typical error was calculated as the standard deviation in each subject's measurements between tests, after any shifts in the mean had been taken into account, and was expressed as a percentage of the subject's mean score to obtain a more easily interpretable percentage score. This percentage is also known as the "coefficient of variation." Total error was calculated as the average of all individual standard deviations for heart rate and time, based on the data of the 2 trials.[34] The level of statistical significance was set at $p = .05$.

In order to assess the amount of error associated with repeated measurements, the standard error of measurement (SEM) was calculated.[34] Standard errors of measurement between the 2 shuttle run test sessions were computed applying a 95% confidence interval. To determine whether there were significant differences for peak heart rate and exercise time between the treadmill tests and the shuttle run tests, the data were compared using paired *t* tests. Correlation coefficients (Pearson *r*) for VO_{2peak} achieved during treadmill tests and shuttle run tests were computed for validity. A linear regression analysis with backward elimination procedure also was performed to determine which measures could significantly predict VO_{2peak} in children and adolescents with CP.

Results

Peak RER was ≥1.0 and peak heart rate was >180 bpm in all subjects. These variables indicate that a maximal effort was reached during all tests.

Test-Retest Reliability for 10-m Shuttle Run Tests

The physiological variables measured during both shuttle run tests are described in Table 2. The test-retest reliability statistics of both exercise performances are shown in Table 3 and illustrated in Figures 1 and 2. Intraclass correlation coefficients (2-way mixed) for heart rate and exercise time were .87 or above for both shuttle run tests.

There were no significant differences in maximal heart rate and total exercise time between the 2 shuttle run tests. The results indicate that both 10-m shuttle run tests have good reproducibility.

The SEM values are shown in Table 2. The SEM values for exercise time ranged from 0.25 for the SRT-II to 0.42 for the SRT-I, and the SEM values for peak heart rate ranged from 1.52 for the SRT-II to 2.56 for the SRT-I.

Validity for 10-m Shuttle Run Tests and Treadmill Tests

Both 10-m shuttle run tests were compared with the treadmill tests for validity of VO_2 values. The peak heart rate values obtained during the GMFCS level I treadmill test ($\overline{X}$=192.9 bpm, SD=6.2) and the GMFCS level II treadmill test ($\overline{X}$=193.1 bpm,

Table 2. Reproducibility (Test-Retest) of Gross Motor Function Classification System (GMFCS) Level I and II Shuttle Run Test Measurements. (n=25)[a]

	Measurement 1		Measurement 2			
	Mean	SD	Mean	SD	Change in Mean	SEM
GMFCS level I						
HR_{peak} (bpm)	200.6	6.6	198.9	6.6	1.7	2.56
Exercise time (min)	8.2	2.3	8.6	2.0	0.4	0.42
GMFCS level II						
HR_{peak} (bpm)	197.9	6.6	197.3	5.2	0.6	1.52
Exercise time (min)	11.5	3.8	11.6	3.8	0.1	0.25

[a]Change in mean denotes the change between measurement 1 and measurement 2; SEM=standard error of measurement; HR_{peak}=peak heart rate.

Table 3. Reliability (Test-Retest) Statistics of the Gross Motor Function Classification System (GMFCS) Level I and II Shuttle Run Tests. (n=25)[a]

	ICC	Typical Error	Total Error	LOA	Typical Error % (CV)
GMFCS level I					
HR_{peak} (bpm)	.87	2.40	2.65	6.63	1.2
Time (min)	.97	0.36	0.43	1.01	4.8
GMFCS level II					
HR_{peak} (bpm)	.94	1.56	1.52	4.32	0.8
Time (min)	.99	0.31	0.30	0.86	2.7

[a]ICC=intraclass correlation coefficient, LOA=limits of agreement, CV=coefficient of variation, HR_{peak}=peak heart rate.

SD=6.1) were significantly lower ($p < .05$) from those obtained during the GMFCS level I and II shuttle run tests ($\bar{X}$=200.6 bpm, SD=6.7, and $\bar{X}$=199.4 bpm, SD=6.8, respectively). The subjects reached this higher peak heart rate in a significantly ($p < .05$) shorter time during the shuttle run tests compared with the treadmill tests (Tab. 4).
The physiological variables measured on both tests in which VO_{2peak} values were obtained are described in Table 4. Validity statistics for the shuttle run tests and the treadmill test in which VO_{2peak} values were obtained are shown in Table 5 and Figure 3. Pearson correlation coefficients for VO_{2peak} achieved during the shuttle run test with gas analysis and the treadmill test were .96 for subjects with a level I classification on the GMFCS and .96 for subjects with a level II classification on the GMFCS. The results indicate that both 10-m shuttle run tests are valid measures of aerobic capacity (VO_2) in children and adolescents with CP.

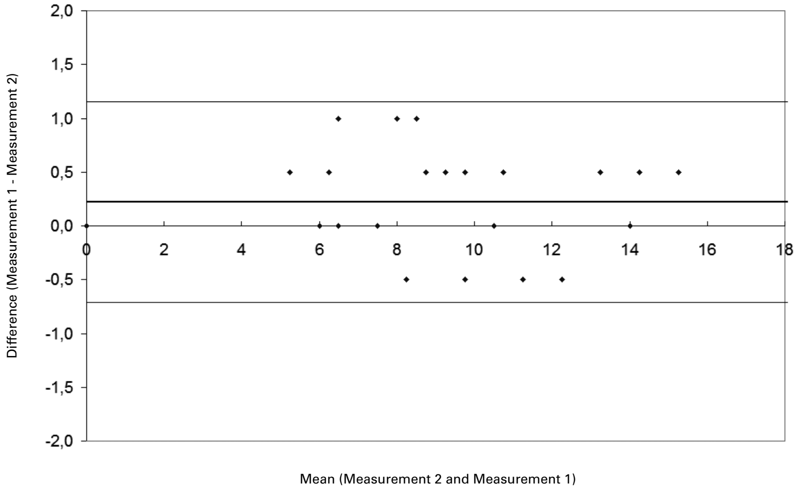

Figure 1. Bland-Altman plot of exercise time during measurement 1 and measurement 2 for Gross Motor Function Classification System (GMFCS) Levels I and II. The bold-type line shows the mean difference between the 2 measurements, and the 2 thin lines indicate ±2 standard deviations. On the X-axes, the average exercise time from both tests is displayed. On the Y-axes, the difference between the exercise times during measurement 1 and measurement 2 is displayed.

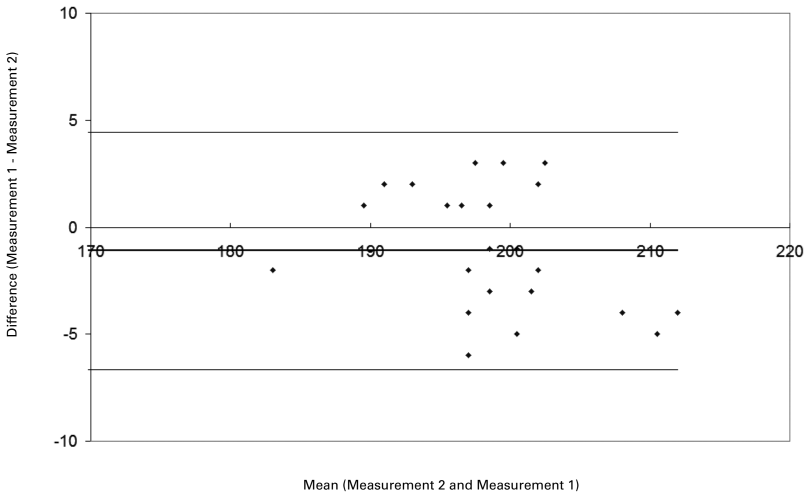

Figure 2. Bland-Altman plot of peak heart rate (HR_{peak}) during measurement 1 and measurement 2 for Gross Motor Function Classification System (GMFCS) Levels I and II. The bold-type line shows the mean difference between the 2 measurements, and the 2 thin lines indicate ±2 standard deviations. On the X-axes, the average HR_{peak} from both tests is displayed. On the Y-axes, the difference between the HR_{peak} during measurement 1 and the HR_{peak} during measurement 2 is displayed.

Table 4. Comparison of Gross Motor Function Classification System (GMFCS) Level I and II Shuttle Run Tests With Gas Analysis and Treadmill Tests. (n=24)[a]

	10-m SRT-STG		**Treadmill Tests**		
	Mean	**SD**	**Mean**	**SD**	**Change in Mean**
GMFCS level I					
VO_{2peak} (L/min)	1.7	0.5	1.7	0.5	0
RER	1.1	0.1	0.9	0.9	0.2
HR_{peak} (bpm)[b]	200.6	6.7	192.9	6.2	7.7
Exercise time (min)[b]	8.2	2.2	10.6	4.0	2.4
GMFCS level II					
VO_{2peak} (L/min)	1.7	0.6	1.6	0.6	0.1
RER	1.0	0.1	1.0	0.8	0
HR_{peak} (bpm)[b]	199.4	6.8	193.1	6.1	6.3
Exercise time (min)[b]	11.5	3.8	13.4	4.1	1.9

[a]SRT-STG=shuttle run tests with gas analysis, VO_{2peak}=peak oxygen uptake, RER=respiratory exchange ratio, HR_{peak}=peak heart rate. In one subject, VO_{2peak} and RER were not monitored, and therefore this subject's data were not included in the validity analysis.

[b]$p < .05$ for the comparison of the 10-m SRT and the treadmill test.

Table 5. Validity Statistics[a] of the Shuttle Run Tests With Gas Analysis and the Treadmill Tests (n=24)[b]

	r	**Typical Error**	**95% CI of Typical Error**	**Total Error**	**LOA**	**Typical Error % (CV)**
GMFCS level I						
VO_{2peak} (L/min)	.96	0.10	0.07-0.17	0.10	0.28	5.4
GMFCS level II						
VO_{2peak} (L/min)	.96	0.13	0.09-0.23	0.15	0.37	7.1

[a]Pearson product moment correlation (r).

[b]CI=confidence interval, LOA=limits of agreement, CV=coefficient of variation, VO_{2peak}=peak oxygen uptake. In one subject, VO_{2peak} and RER were not monitored, and therefore this subject's data were not included in the validity analysis.

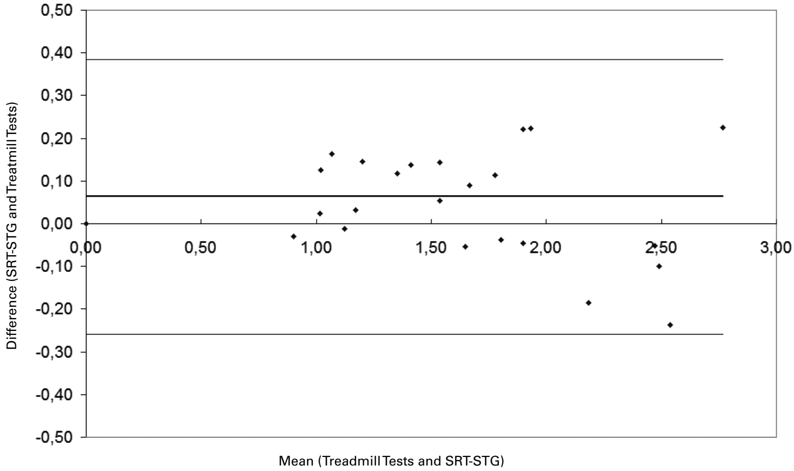

Figure 3. Bland-Altman plot of peak oxygen uptake (VO_{2peak}) during shuttle run tests with gas analysis STG) and treadmill tests for Gross Motor Function Classification System (GMFCS) Levels I and II. The bold-type line shows the mean difference between the 2 measurements, and the 2 thin lines indicate ±2 standard deviations. On the X-axes, the average VO_{2peak} from both tests is displayed. On the Y-axes, the difference between the VO_{2peak} during the treadmill tests and the VO_{2peak} during the STG is displayed.

Feasibility

The tests were generally well tolerated by the subjects. At the completion of tests, subjects were asked a standardized question: "Which of the 2 tests, shuttle run or treadmill, did you prefer, and why?" Of the 25 subjects, 23 preferred the shuttle run test over the treadmill test, 1 preferred the treadmill test over the shuttle run test, and 1 did not favor any type of testing. Predominant reasons for the shuttle run test preference were that the subjects felt they were in control of the test and did not need to acquire new skills. Several subjects added that they could "try harder" with the shuttle run test than in the treadmill test because they knew they "could stop at any time." Moreover, their gait disturbance was less of a problem during this field exercise test compared with the treadmill test.

A backward regression analysis indicated that VO_{2peak} or VO_{2peak}/kg could not be predicted from treadmill exercise time, maximal treadmill running speed, maximal shuttle run test running speed, or GMFCS level. The only significant predictors of VO_{2peak} and VO_{2peak}/kg were sex and body weight: $\overline{X}=0.29 \times \text{weight} - 0.474(\text{sex})$[1=male, 2=female] + 1,159; r=.84, $p < .0001$; standard estimate of error=0.31.

Discussions and conclusions

Field testing is widely used to estimate aerobic capacity in children and adolescents because it is impractical to obtain laboratory measurements for large groups of people.[14,35-38] Although it is generally agreed that cardiovascular field tests are good measures of cardiorespiratory fitness, it is sometimes difficult to compare individuals' performance on different field tests because scores are not often converted into VO_{2peak} values. Furthermore, to date, there are no validated field tests for children and adolescents with CP. The results of this study demonstrate that both recently developed 10-m shuttle run tests are reproducible, clinical applicable, and valid measures in children and adolescents with CP.

The shuttle run tests and treadmill tests were similar in that steady increments in heart rate with an increase in exercise intensity were apparent. The peak heart rate during the treadmill tests was different from that obtained during the shuttle run tests. The subjects reached a higher peak heart rate in a shorter time during the shuttle run tests. As with conventional maximal tests, there was a linear relationship between heart rate and VO_2 in the shuttle run and treadmill tests. Thus, while the newly developed shuttle run tests do not in any way challenge the use and significant benefits of conventional methods of exercise testing, our study suggests that both shuttle run tests have several characteristics that are similar to the accepted gold standard.

A limitation of this study is the fact that the treadmill VO_2 setup has never been validated in children and adolescents with CP. However, pilot testing showed that measurements of maximal VO_2 during an exercise test on the treadmill were reproducible. Together with the high maximal heart rates obtained in our subjects and the high RER values,[39] one might assume that treadmill testing yields valid measurements of maximal VO_2 in children and adolescents with CP. Therefore, we used the treadmill protocol as the gold standard, and we compared the results of the treadmill tests with the results of both shuttle run tests.

There are important differences between the shuttle run tests and the treadmill tests. First, during the shuttle run tests, a child is allowed to run faster than the set speed. The energy cost of changing direction (by 180° in the case of the shuttle run tests) and of decelerating and then accelerating is greater than continually running in a straight line on a treadmill. Children who are efficient in changing direction will be favored in this test compared with children who are poor in changing direction. However, in daily life, children require repeated changes of direction. Thus, the shuttle run tests may be more useful in predicting physical performance than a continuous gas analysis VO_{2peak} test on a treadmill. Second, when running on a treadmill, it is not possible to run faster or slower than the treadmill, because the running speed is externally paced. While performing the shuttle run tests, it is possible for a person to walk faster than the speed determined by the signal.

For some children, it is difficult to run at the audio signal that determines the running speed. Therefore, it is recommended that during the first stages of the test, a person should assist the child.[40] If the child understands the principle of the shuttle run tests, he or she can continue the test without assistance. If children have problems pacing themselves, they should be accompanied throughout the test. In these situations, an extra person to collect reliable and valid shuttle run test information is necessary.

In our study, there was no difference between the shuttle run tests for heart rate and exercise time. However, in both 10-m shuttle run tests, the exercise time was shorter than in the treadmill tests. This difference may have been due to the fact that a lot of energy was used to make a turn every 10 m. Some authors[10,41] have stated that the energy used for postural stabilization, high muscle tone, and involuntary movements causes an increase in VO_2 during submaximal exercise. During the treadmill tests in our study, the energy used for stabilization may have been low because the subjects were allowed to hold on to the guardrails with their fingers.

In the rehabilitation and fitness setting, VO_{2peak} is often estimated from the maximal speed and grade attained during a maximal treadmill exercise test.[28] For the exact determination of VO_{2peak} during exercise, a respiratory gas analysis system is necessary. However, this equipment is expensive and not always available. It would be of value if the VO_{2peak} could be predicted from surrogate measures in children with CP, because many clinicians do not have a respiratory gas analysis system. Unfortunately, in our study, neither maximal running speed nor treadmill exercise time could accurately predict VO_{2peak} in children with CP. Therefore, the shuttle run tests could be used only as an instrument to monitor changes in exercise capacity. The exercise time or level achieved is the most useful outcome measure. The heart rate during the shuttle run tests can be used to check whether a child has performed maximally (heart rate >180 bpm).

The calculated SEM can be used to determine the range in which a person's "true score" could be expected to lie, considering the amount of error associated with repeated measures. For example, we can be 95% confident that the "true score" for people performing the shuttle run tests lies within ±2 SEM. Thus, a change in person's performance of greater than 2 SEM most likely represents a real change that may not be attributed to measurement error. Based on the data in the "Results" section, total increases of >0.84 (2 × 0.42) minute for the SRT-I and 0.50 (2 × 0.25) minute for SRT-II could be attributed to real change with 95% confidence.

In this study, the subjects demonstrated a preference for the 10-m shuttle run tests over the treadmill tests. The shuttle run tests are nonthreatening and can easily be performed. A person can terminate the test at any point that he or she chooses. Moreover, as shuttle run tests require a child to either run or walk between 2 markers, the tests do not necessitate acquisition of new skills to participate. Shuttle run tests can be widely used in a mainstream school, schools for special education, or rehabilitation centers. None of the children or adolescents in our study required extensive instructions to participate. Both field tests seem to be useful for evaluating the aerobic performance of children and adolescents who are classified at level I or II on the GMFCS.

References

1. Unnithan VB, Clifford C, Bar-Or O. Evaluation by exercise testing of the child with cerebral palsy. Sports Med. 1998;26:239-251.
2. Fernhall B, Millar AL, Tymeson GT. Maximal exercise testing of mentally retarded adolescents and adults: reliability study. Arch Phys Med Rehabil. 1990;71:1065-1068.
3. Astrand PO, Rodahl K. Textbook of Work Physiology: Physiological Bases of Exercise. 4^{th} ed. Champaign, Ill: Human Kinetics Inc; 2003.
4. Lundberg A. Longitudinal study of physical working capacity of young people with spastic cerebral palsy. Dev Med Child Neurol. 1984;26:328-334.
5. Bar-Or O, Inbar O, Spira R. Physiological effects of a sports rehabilitation program on cerebral palsied and post-poliomyelitic adolescents. Med Sci Sports. 1976;8:157-161.
6. Shephard RJ. Tests of maximum oxygen intake: a critical review. Sports Med. 1 984;1:99-124.
7. Turley KR, Rogers DM, Wilmore JH. Maximal testing in prepubescent children: treadmill versus cycle ergometry. Med Sci Sports. 1993;25:S9.
8. Armstrong N, van Mechelen W. Pediatric Exercise Science and Medicine. New York, NY: Oxford University Press Inc; 2000.
9. Wiley ME, Damiano DL. Lower-extremity strength profiles in spastic cerebral palsy. Dev Med Child Neurol. 1998;40:100-107.
10. Lundberg A. Maximal aerobic capacity of young people with spastic cerebral palsy. Dev Med Child Neurol. 1978;20:205-210.
11. Berg-Emons van den RJ, van Baak MA, de Barbanson DC, et al. Reliability of tests to determine peak aerobic power, anaerobic power and isokinetic muscle strength in children with cerebral palsy. Dev Med Child Neurol. 1996;38:1117-1125.
12. Hoofwijk M, Unnithan VB, Bar-Or O. Maximal treadmill performance of children with cerebral palsy. Pediatr Exerc Sci. 1995;7:305-313.
13. Massin M, Allington N. Role of exercise testing in the functional assessment of cerebral palsy children after botulinum toxin injection. J Pediatr Orthop. 1999;19:362-365.
14. Leger LA, Mercier D, Gadoury C, Lambert J. The multistage 20 meter shuttle test for aerobic fitness. J Sports Sci. 1988;6:93-101.
15. Leger LA, Lambert J. A maximal multistage 20-m shuttle run test to predict VO2max. Eur J Appl Physiol Occup Physiol. 1982;49:1-12.
16. Boreham CAG, Palicka VJ, Nichols AK. A comparison of the PWC170 and 20-MST tests of aerobic fitness in adolescent schoolchildren. J Sports Med Phys Fitness. 1990;30:19-23.
17. Liu NYS, Plowman SA, Looney MA. The reliability and validity of the 20 m shuttle run test in American students 12 to 15 years old. Res Q Exerc Sport. 1992;63:360-365.
18. Mechelen WV, Hlobil H, Kemper HCG. Validation of two running tests as estimates of maximal aerobic power in children. Eur J Appl Physiol. 1986;55:503-506.
19. Ramsbottom R, Brewer J, Williams C. An estimation of the maximal multistage shuttle run test as a predictor of VO2max in active female subjects. J Sports Sci. 1988;6:165.

20. Paliczka VJ, Nichols AK, Boreham CAG. A multistage shuttle run as a predictor of running performance and maximal oxygen uptake in adults. Br J Sports Med. 1987;21:163-165.
21. Buchfuhrer MJ, Hansen JE, Robinson TE, et al. Optimizing the exercise protocol for cardiopulmonary assessment. J Appl Physiol. 1983;55:1558-1564.
22. Orenstein DM. Exercise testing in cystic fibrosis. Pediatr Pulmonol. 1998;35:223-225.
23. Palisano RJ, Rosenbaum P, Walter S. The development and reliability of a system to classify gross motor function in children with cerebral palsy. Dev Med Child Neurol. 1997;39:214-223.
24. Pollack ML, Schmidt DH, Jackson AS. Measurement of cardio-respiratory fitness and body composition in the clinical setting. Comp Ther. 1980;6:12-27.
25. Armstrong LE, Whaley MH, Brubaker PH, Otto RM. ACSM's Guidelines for Exercise Testing and Prescription. 7th ed: Philadelphia, Pa: Lippincott Williams & Wilkins; 2005.
26. MacFarlane DJ, Fogarty BA, Hopkins WC. The accuracy and variability of commercially available heart rate monitors. New Zealand J Sports Med. 1989;17:51-53.
27. Medbo JL, Mamen A, Welde B, et al. Examination of the Metamax I and II oxygen analysers during exercise studies in the laboratory. Scand J Clin Lab Invest. 2002;62:585-598.
28. Meyer T, Georg T, Becker C, Kindermann W. Reliability of gas exchange measurements from two different spiroergometry systems. Int J Sports Med. 2001;22:593-597.
29. Brehm MA, Harlaar J, Groepenhof H. Validation of the portable VmaxST system for oxygen-uptake measurement. Gait Posture. 2004;20:67-73.
30. Jones AM, Doust JH. A 1% treadmill grade most accurately reflects the energetic cost of outdoor running. J Sports Sci. 1996;14:321-327.
31. Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice. East Norwalk, Conn: Appleton & Lange; 2000.
32. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1:307-310.
33. Hopkins WG. Measures of reliability in sports medicine and science. Sports Med. 2000;30:1-15.
34. Knapp TR. Technical error of measurement: a methodological critique. Am J Phys Anthropol. 1992;87:235-236.
35. Fernhall B, Pitetti KH, Stubbs N, Stadler L. Validity and reliability of the ½ mile run-walk as an indicator of aerobic fitness in children with mental retardation. Pediatr Exerc Sci. 1996;8:130-142.
36. Selvadurai HC, Cooper PJ, Meyers N. Validation of shuttle tests in children with cystic fibrosis. Pediatr Pulmonol. 2003;35:133-138.
37. Cureton KJ, Sloniger MA, O'Bannon JP, et al. A generalized equation for prediction of VO_{2peak} from 1-mile run-walk performance. Med Sci Sports Exerc. 1995;27:445-451.
38. Cooper SM, Baker JS, Tong RJ, et al. The repeatability and criterion related validity of the 20 m multistage fitness test as a predictor of maximal oxygen uptake in active young men. Br J Sports Med. 2005;39:19-24.
39. Howley ET, Bassett DRJ, Welch HG. Criteria for maximal oxygen uptake: review and commentary. Med Sci Sports Exerc. 1995;27:1292-1301.

40. Guerra M, Pitetti KH, Fernhall B. Cross validation of the 20-meter shuttle run test for adolescents with Down syndrome. Adapt Phys Act Q. 2003;20:70-79.

41. Lundberg A. Oxygen consumption in relation to work load in students with cerebral palsy. J Appl Physiol. 1976;40:873-875.


IN CHILDREN AND ADOLESCENTS
00.06.45
00.51.02
WITH CEREBRAL PALSY