Test–retest reliability of the 10-metre fast walk test and 6-minute walk test in ambulatory school-aged children with cerebral palsy

Patricia Thompson HBSc(Kin) MScPT;
Tricia Beath BScH(Life Sci) MScPT;
Jacqueline Bell BScH(Life Sci) MScPT;
Gabrielle Jacobson BSc(Biol) MScPT;
Tegan Phair BScH MScPT;
Nancy M Salbach BSc(Physiol) BSc(PT) MSc PhD, Lecturer and Postdoctoral Fellow;
F Virginia Wright* BSc(PT) MSc PhD, Clinician Scientist, Bloorview Research Institute, Toronto, Canada.

**Correspondence to last author at* Bloorview Research Institute/Bloorview Kids Rehab, 150 Kilgour Road, Toronto, ON, M4G 1R8, Canada.
E-mail: vwright@bloorview.ca

DOI:10.1111/j.1469-8749.2008.02048.x
Published online 18th March 2008

Short-term test–retest reliability of the 10-metre fast walk test (10mFWT) and 6-minute walk test (6MWT) was evaluated in 31 ambulatory children with cerebral palsy (CP), with subgroup analyses in Gross Motor Function Classification System (GMFCS) Levels I (*n*=9), II (*n*=8), and III (*n*=14). Sixteen females and 15 males participated, mean age 9 years 5 months (SD 3y 7mo, range 4y 3mo–18y 2mo). Twenty had spastic diplegia, while the others had another form of CP. Retest interval varied from 1 to 4 weeks (mean 10.6d [SD 6.4]). Intraclass correlation coefficients (ICCs) estimated reliability. The 10mFWT ICC was 0.81 (95% confidence interval [CI] 0.65–0.90) across participants, and >0.59 in GMFCS subgroups (95% CI lower bound >0.01). The 6MWT ICC was 0.98, and >0.90 in GMFCS subgroups (95% CI lower bound >0.64). Bland–Altman plots indicated bias towards higher 6MWT retest distances in GMFCS Level I. Minimum detectable change (95% CI) was 61.9, 64.0, and 47.4m for the 6MWT within GMFCS Levels I, II, and III respectively. The conclusion is that while the 10mFWT showed inadequate test–retest reliability given its wide 95% CI, the 6MWT demonstrated good to excellent reliability. Investigation of the need for a practice walk when administering the 6MWT with children in GMFCS Level I is recommended to establish their fastest pace.

See end of paper for list of abbreviations.

Cerebral palsy (CP) is a non-progressive motor impairment caused by an irreversible insult to the nervous system. The motor deficits have detrimental effects on speed,[1] distance parameters,[2] and quality of gait. Consequently, children with CP frequently engage in therapy directed at improving walking skills so they can participate more fully in activities of daily living (ADL) and recreational activities.

The 10-metre fast walk test (10mFWT) and 6-minute walk test (6MWT) have the potential to provide valuable clinical information regarding gait abilities and outcomes in these children. Both are safe, easy, and inexpensive to administer.[3,4] The 10mFWT, which represents the minimum distance required for functional ambulation,[3] has demonstrated excellent test–retest and interrater reliability in adults with traumatic brain injury (TBI),[3,4] and high interrater reliability with healthy participants.[3]

The 6MWT is a self-paced, submaximal test that assesses functional capacity for walking a prolonged distance.[4,5] It may reflect exercise tolerance required for the performance of ADL,[5] and predict ability to walk in the community.[4] Its psychometric properties have been examined in various populations.[4–17] Excellent test–retest reliability has been observed among adults with TBI[4,10] and multiple sclerosis (MS).[11] It has also been evaluated in healthy children,[12] children with juvenile idiopathic arthritis,[13,14] and children awaiting organ transplant[15] with indication of psychometric acceptability in these groups.

Despite the potential applicability of the 10mFWT and 6MWT for children with CP, to our knowledge, there are no published reports on reliability with this group. Thus, the primary objective of this study was to estimate the test–retest reliability of the 10mFWT and 6MWT in ambulatory school-aged children with CP and within Gross Motor Function Classification System (GMFCS)[18] subgroups (i.e. GMFCS Levels I, II, and III). The secondary objective was to estimate the minimum detectable change (MDC) in the total sample and subgroups.

Method

A longitudinal pilot study with a baseline and retest session was conducted. The retest interval was 1 to 2 weeks, a period during which changes in measures of walking speed and distance were not expected. A maximum 4-week retest interval was allowed.

PARTICIPANTS

Participants met the following inclusion criteria: (1) diagnosis of spastic CP; (2) between 4 and 18 years of age; (3) GMFCS Levels I, II, or III as determined by their developmental pediatrician; (4) ability to walk independently without stopping for 6 minutes, with or without a walking aid; (5) ability to follow verbal instructions in English; and (6) ability to cooperate for at least 30 minutes as judged by their treating physiotherapist (PT). The exclusion criteria were: (1) orthopaedic surgery within the past 6 months; and (2) botulinum toxin type A (BoNT-A) injections within the preceding 3 months. Participants were enrolled from a convenience sample of children attending physiotherapy or an integrated education and therapy school program at our children's treatment centre. Ethics approval was granted from the University of Toronto and Bloorview Kids Rehab (BKR). Written informed consent to participate in the study was obtained from a legal guardian of each participant, and written assent was also given by children over

7 years of age.

DATA COLLECTION

The investigative team comprised two PTs from BKR and five PT students. The investigators paired themselves into six assessor teams to provide a flexible testing schedule for families. In five of the teams, the senior author (FVW) served as the PT doing the timing. Whenever possible, the same assessor team evaluated a participant at both test sessions, and the assessor assumed the same role. A separate data collection form was used for each session to eliminate scoring bias from knowledge of prior results. The type of gait aid, orthosis, and footwear worn at each test session was recorded. Participants were reminded to bring these for retesting. Assessors noted any testing issues such as difficulty in understanding the instructions, loss of focus during testing, need for additional encouragement, difficulties with a walking device, as well as previous experience performing the 6MWT.

10MFWT

The start of the 10m course was demarcated with a tape line on the floor in the mid-stretch of a 30m corridor. Participants positioned their toes behind the start line and were instructed to walk at their fastest speed and continue down the corridor until told to stop. The stop command was given approximately 5m past the finish line so that children would not decelerate until after the 10m mark.[3] A vertical line taped on the adjacent wall demarcated the finish line, as a line on the floor might have encouraged children to alter their speed.

One assessor stood at the start line and instructed the child to start walking when ready,[3] and also encouraged them to keep going as fast as possible. The other assessor stood at the finish line opposite the wall marker and operated the digital stopwatch, recording the time taken to walk 10m (to the nearest tenth of a second). Timing began the moment the participant initiated a step[3] and ended when the leading foot crossed the finish line.

Participants were given a practice trial to ensure adequate understanding of the instructions prior to performing two timed trials. The assessor first demonstrated what a fast walk and what a run looked like. If a child ran during the practice test, they were called back and a second demonstration and practice were performed to correct this. Two test trials were then performed. Participants received a 1 to 5-minute seated rest prior to the 6MWT, starting again when ready. The identical protocol, including practice test, was repeated at the 10mFWT retest.

6MWT

The walking course was set in a quiet, rectangular corridor (20m x 45m) and marks were taped on the walls at 30m intervals. Children were instructed to walk as many laps as possible in 6 minutes without running. A practice trial was not given for this lengthy test because of concerns about the impact of associated fatigue[5,11] on the subsequent test, and in light of evidence[4] that a familiarization trial does not enhance reliability.

One assessor walked just behind the participant to provide standardized encouragement every 30s (e.g. 'great work' and 'keep it up') as recommended in the American Thoracic Society (ATS) 6MWT guidelines.[5] The second assessor walked approximately 1m behind the participant and monitored the stopwatch. Timing began the moment the participant initiated a step. At 6 minutes, the trailing PT marked the participant's finish point on the wall. The cumulative distance to the nearest 30m mark was computed and any remaining distance walked was measured to the nearest centimetre. The same protocol was repeated at retest. When

Table I: Participant characteristics

	All participants (n=31)	*GMFCS Levels I, II (n=17)*	*GMFCS Level III (n=14)*
Sex			
Male	15	8	7
Female	16	9	7
Age, y			
4–6	8	4	4
7–10	10	7	3
11–14	9	4	5
15–18	4	2	2
Mean (SD)	9y 5mo (3y 7mo)	9y 2mo (3y 6mo)	9y 7mo (4y)
range	4y 3mo–18y 2mo	5y 3mo–18y 2mo	4y 3mo–16y 6mo
Diagnosis			
Spastic diplegia	20	10	10
Spastic triplegia	2	1	1
Spastic quadriplegia	2	0	2
Hemiplegia	4	4	0
Other	3	2	1
Time between tests			
≤14d	26	12	14
>14d	5	5	0
Mean (SD)	10.6 (6.4)	12.2 (7.9)	8.6 (2.5)
range	7.0–31.0	7.0–31.0	7.0–14.0

GMFCS, Gross Motor Function Classification System.

younger participants had difficulty understanding the concept of walking continuously for 6 minutes, the assessors provided a series of visual goals approximately 20m ahead of the participant throughout the test. Each time a participant passed a visual goal, a new goal was immediately targeted 20m further along.

STATISTICAL ANALYSIS

Analyses were conducted using the Statistical Package for Social Sciences (SPSS) version 15.0. Intraclass correlation coefficients (ICC; type 2:1)[19] and associated 95% confidence intervals (CI) estimated test–retest reliability. An ICC ≥0.80 reflects excellent reliability, while ICCs from 0.70 to 0.79 reflect good reliability.[19] The recommended minimum for the lower bound of the 95% CI is 0.85.[20] The10mFWT ICCs were computed using the second trial for both baseline and retest as participants were expected to understand maximally the test requirements by then.

In connection with the ICC, standard error of measurement (SEM) was computed such that $SEM = SD_b*(\sqrt{1-ICC})$, where SDb is the SD of baseline scores.[21] The MDC of each measure was calculated at a 95% CI to provide clinicians with information to determine whether scores on repeat evaluation reflect true change, where $MDC95\ (95\%\ CI) = 1.96*\sqrt{2}*SEM$.[19]

Paired *t*-tests were used to determine differences between GMFCS levels within each test period and to assess changes in scores from test to retest.[22] *p* values were adjusted to 0.01 to handle multiple comparisons.[19]

The Bland–Altman method[23] evaluated measurement bias. This method plots individual difference scores from test and retest against the mean difference score. The coefficient of variation (CV) and limits of agreement (LOA) were also estimated. For the CV, the typical error was first calculated as the SD of the test–retest difference scores such that method error (ME) = $SD_{diff}/v2$. The $CV_{ME} = 2ME/(X_1–X_2) \times 100$, where X_1 and X_2 are the mean group scores for test and retest.[19] The arbitrarily chosen goal for a CV is 10% or below.[22] The LOA was calculated as the mean test–retest difference [d] +/- $(1.96 \times SD_{diff})$.[23] The difference between test and retest for any individual should fall 95% of the time within the interval between the estimated upper and lower limits.[22]

Analyses were completed with all participants and within GMFCS Levels I, II, and III. ICCs were also computed for each walk test for children with and without identified testing issues, and then for subgroups defined by previous experience performing the 6MWT in therapy sessions.

Results

PARTICIPANT CHARACTERISTICS

Thirty-two participants were tested at baseline. One participant in GMFCS Level III did not return because of transportation issues; this child was not considered in the analysis. One other participant (in GMFCS Level III) was able to complete the 10mFWT at each session but not the 6MWT because of difficulties in focusing on the task. Table I presents participant characteristics at baseline for the entire sample and by GMFCS level. Twenty participants had spastic diplegic CP, while the others had another form of CP. Participants' ages varied from 4 years 3 months to 18 years 2 months (mean age 9y 5mo). The mean ages of children in GMFCS Levels I, II, and III were comparable. All but five participants were retested within the targeted 14-day interval (mean interval 10.6d [SD 6.4]). Nine children were reassessed by a team in which one of the two assessors was a different person from baseline.

Table II: 10-metre fast walk test (s)

	Baseline			*Retest*		
	n	*Mean (SD)*	*Range*	*n*	*Mean (SD)*	*Range*
All participants	31	11.4 (10.2)	4.4–57.9	31	9.9 (5.6)	4.4–31.7
GMFCS Level I	9	5.9 (1.0)[a,b]	4.4–7.5	9	6.5 (0.7)[c,d]	5.3–7.3
GMFCS Level II	8	9.6 (2.8)[a]	6.9–15.1	8	8.7 (2.6)[c]	5.6–12.8
GMFCS Level III	14	16.0(13.7)[b]	4.5–57.9	14	12.9 (7.2)[d]	4.4–31.7

[a,b,c,d] Significant differences (p<0.015) between Gross Motor Function Classification System (GMFCS) levels are denoted for the respective comparisons. Pairwise comparisons (p value adjusted to 0.01 for multiple comparisons) for differences between GMFCS levels, and between test and retest scores. There was no difference for any of the test–retest comparisons within GMFCS levels although these were limited by study power (maximum power <0.40).

Table III: Six-minute walk test (m)

	Baseline			*Retest*		
	n	*Mean (SD)*	*Range*	*n*	*Mean (SD)*	*Range*
All participants	30	333.5 (144.8)	81.5–662.0	30	340.8 (155.0)	106.6–669.0
GMFCS Level I	9	486.6 (84.4)[a,b]	385.3–662.0	9	504.2 (102.1)[c,d]	392.7–662.0
GMFCS Level II	8	312.9 (77.0)[a]	193.0–424. 7	8	296.9 (103.3)[c]	111.7–452.8
GMFCS Level III	13	240.2 (121.1)[b]	81.5–429.8	13	254.7 (124.8)[d]	106.6–451.0

[a,b,c,d] Significant differences (p<0.015) between Gross Motor Function Classification System (GMFCS) levels are denoted for the respective comparisons. Pairwise comparisons (p value adjusted to 0.01 for multiple comparisons) for differences between GMFCS levels, and between test and retest scores. There was no significant difference for any of the test–retest comparisons within GMFCS levels although these were limited by study power (maximum power <0.30).

10MFWT TEST–RETEST RELIABILITY

Mean walking times across all participants at baseline and retest were 11.4s and 9.9s respectively (Table II); this difference was not significant ($p=0.14$). There was no significant difference between test and retest within GMFCS levels (Table II), although the power of all comparisons was less than 0.40. The 10mFWT distinguished between children in GMFCS levels with baseline mean times varying from 5.9s in Level I to 16s in Level III ($p<0.015$; Table II).

The test–retest ICC across all participants was 0.81 (95% CI 0.65–0.90) with an SEM of 4.4s and MDC_{95} of 12.2s (Table IV). GMFCS level ICCs are provided in Table IV. The ICC estimates and lower bound of the 95% CI appeared slightly lower when there were different assessors at retest compared with use of the same assessor team, and among children identified as having testing issues compared with those with no testing issues (Table V).

The CVs were in the 16 to 40% range for all analyses (Table IV) with the exception of children in GMFCS Level I for whom the CV was 7.9%. Issues with LOA were most evident for GMFCS Level III, in which an individual's retest score could be as much as 18.8s above or 12.2s below the initial value. This wide interval is in line with the MDC of 17.7s in Level III.

In the Bland–Altman plot (Fig. 1), there is weak evidence of measurement bias such that those with slower walk times at baseline, i.e. children in GMFCS Level III with baseline scores above 15s, had faster walk times on retest. There is an outlier associated with a child in Level III. This child had difficulties with the test, and was the one who could not perform the 6MWT. However, removing this outlier did not change the ICC (i.e. recalculated as 0.78 with 95% CI 0.59–0.89), nor did it strengthen the results for Level III (i.e. recalculated as ICC 0.72 with 95% CI 0.3–0.9). Indeed, the increased variability in scores that resulted from including this child may have inflated the ICC in the total sample.

6MWT TEST–RETEST RELIABILITY

Mean walking distances for the 6MWT across participants

Table IV: Test-retest reliability

	n	*ICC*	*95% CI of ICC*	*Coefficient of variation*	*Limit of agreement (SD)*	*SEM*	*MDC_{95}*
10-metre fast walk test (s)							
All participants	31	0.81	0.65–0.90	36.4	1.5 (10.8)	4.4	12.2
GMFCS Level I	9	0.59	0.01–0.89	7.9	–0.6 (1.3)	0.6	1.7
GMFCS Level II	8	0.70	0.12–0.93	16.0	0.9 (14.0)	1.5	4.3
GMFCS Level III	14	0.78	0.45–0.93	38.5	3.2 (15.4)	6.4	17.7
6-minute walk test (m)							
All participants	30	0.98	0.95–0.99	8.5	–7.3 (64.3)	19.8	54.9
GMFCS Level I	9	0.93	0.71–0.98	4.8	–17.6 (66.0)	22.3	61.9
GMFCS Level II	8	0.91	0.64–0.98	9.0	16.0 (75.6)	23.1	64.0
GMFCS Level III	13	0.98	0.90–1.00	6.1	–14.5 (42.7)	17.1	47.4

GMFCS, Gross Motor Function Classification System; ICC, intraclass correlation coefficient; CI, confidence interval; SEM, standard error of measurement; MDC_{95}, minimal detectable change 95% confidence level.

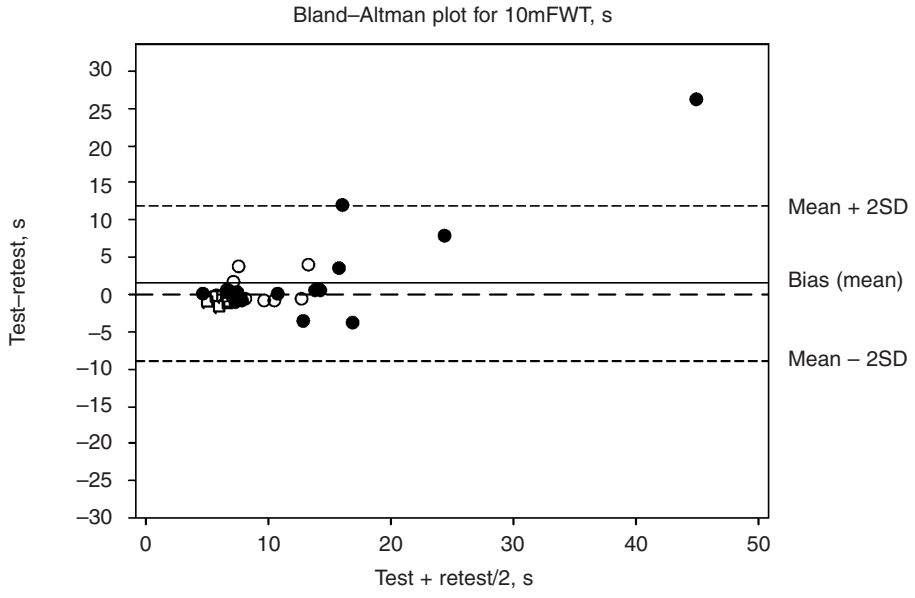

Figure 1: *Bias (mean for test minus retest time) = 1.5s, indicating reduction in time required to cover distance at retest (i.e. increase in speed). 10mFWT, 10-metre fast walk test.*

Black circles, children in Gross Motor Function Classification System (GMFCS) Level III (n=14)
White circles, children in GMFCS Level II (n=8)
Hollow squares, children in GMFCS Level I (n=9).

were 333.5m and 340.8m at baseline and retest respectively (Table III), and the difference was not significant ($p=0.24$). There was no significant difference between test and retest within GMFCS levels either (Table III), although the power of these comparisons was less than 0.30. The test discriminated among children in different GMFCS levels (i.e. baseline mean times varied from 240m in Level III to 486m in Level I; $p<0.015$). The ICC for test–retest reliability across all participants was 0.98 (95% CI: 0.95–0.99) with an SEM of 19.8m and MDC_{95} of 54.9m (Table IV). The lowest ICC was 0.91 in GMFCS Level II.

The CVs were acceptable, all being below 10%. LOA estimates were comparable across subgroups, but were generally 10m to 20m higher than MDC estimates (Table IV). The Bland–Altman plot (Fig. 2) shows evidence of heteroscedastic measurement errors such that children with longer distances at baseline, i.e. those in GMFCS Level I with distances above 450m, increased their walk distances on retest.

Table V: Effect of assessor, familiarity, and test issues on test–retest reliability

	n	*ICC*	*95% CI*
10-metre fast walk test			
Same assessor at retest	22	0.81	0.59–0.91
Different assessor at retest	9	0.75	0.22–0.94
No test issues	23	0.89	0.76–0.95
Test issues	8	0.81	0.32–0.96
6-minute walk test			
No previous experience	14	0.98	0.92–0.99
Previous experience	16	0.99	0.96–1.00
Same assessor at retest	21	0.97	0.93–0.99
Different assessor at retest	9	0.99	0.95–1.00
No test issues	23	0.98	0.95–0.99
Test issues	7	0.96	0.78–0.99

ICC, intraclass correlation coefficient; CI, confidence interval.

Subanalyses that considered use of different assessors on retest and children with testing issues indicated that neither affected the ICCs or 95% CI (Table V), nor was there any impact on ICCs in relation to previous 6MWT experience.

Discussion

The test–retest reliability (ICC=0.81) of the 10mFWT across all participants met the target for acceptability, but fell short of reliability estimates of 10m walking speeds in adults with TBI (i.e. ICC 0.96).[4] The lower bound of the 95% CI was below the 0.85 acceptability standard identified a priori, indicating that the point estimate was too imprecise to support recommendation of this measure. The large CVs also indicated measurement issues.

The limited interparticipant variability in scores in GMFCS Levels I and II (i.e. less than 3s) is likely to have contributed to the subgroups' lower reliability estimates (ICCs<0.70) as compared with that of the more heterogeneous total sample (ICC 0.81). While children in GMFCS Level III had greater between-child variability (SD 13.7s), they also had greater within-participant variability than those in Levels I and II due to the greater influence of spasticity, muscle weakness, and balance issues on their gait. This might point to the need to add a non-timed 2m acceleration phase to the test[9,24] so that the impact of spasticity and stiffness on walking initiation would not be included as a source of measurement error. Paradoxically, the ICC for the Level III group exceeded the 'good reliability' cut-off point of 0.70, perhaps because of the positive impact of the underlying large between-participant variability.[22]

One of the test instructions may also have adversely affected the reliability estimates for the 10mFWT. Specifically, the assessor provided positive feedback on the walk speed that a participant chose when performing the 10mFWT practice test. This was the speed that the child tried to replicate for the walk that was officially timed. When the child returned 2 weeks later and did the practice again, a different 'fastest' walk speed may have been self-selected by the child and reinforced by the assessor. If

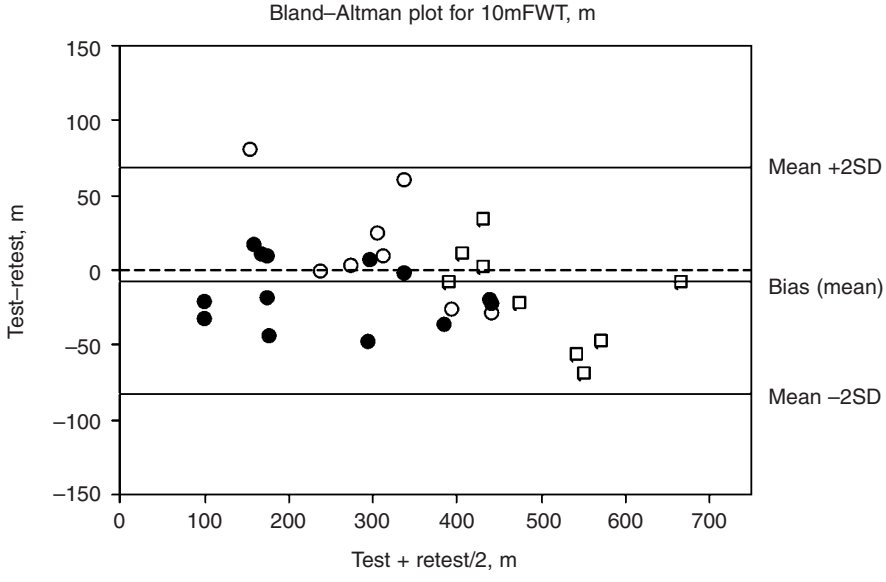

Figure 2: *Bias (mean for test minus retest time)=–7.3m, indicating increase in distance covered at retest. 10mFWT, 10-metre fast walk test.*

Black circles, children in Gross Motor Function Classification System (GMFCS) Level III (n=13)
White circles, children in GMFCS Level II (n=8)
Hollow squares, children in GMFCS Level I (n=9).

children had been asked to run the 10m, they might have had a more consistent performance of their fastest speed. Such issues require further investigation to standardize better the testing protocol.

One other factor that may have led to the lower reliability of the 10mFWT overall was the observed negative impact of having a different assessor at retest for nine of the children. This may have had greatest influence with respect to the inconsistency of assessor reinforcement of the child's selected speed.

In contrast, the 6MWT demonstrated excellent test–retest reliability with a narrow 95% CI in the overall sample and GMFCS subgroups, supporting its use among community ambulators. The ICC estimates are similar to other studies that evaluated the 6MWT test–retest reliability in healthy teenagers,[12] adults with TBI,[4,10] and individuals with MS.[11] Given the lack of evidence for impact of different assessors on reliability, it seems that the 6MWT, with its greater reliance on endurance than achievement of maximal short-burst speed, is more robust to assessor impact than the 10mFWT. The 6MWT CVs were under 10%, lending support to the high ICCs, and were in line with those found by Verschuren et al.[25] in their reliability work with short-distance running tests and muscle tests in children with CP.

Unlike the original 6MWT protocol from the ATS,[5] we did not include a practice trial. Our decision was made based on work by van Loo et al.[4] which demonstrated no significant learning effect when adults with spasticity did two 6MWT trials. There was indication of heteroscedastic measurement errors in GMFCS Level I, i.e. faster children in this group had increased retest distances. These children had perhaps more physical capability to speed up on the retest after perhaps realizing from their first attempt that they could have gone faster. Conversely, children in Level III with their greater physical restriction and associated limits to performance were most consistent in their distance covered as demonstrated within each of the 6MWT analyses.

When using measures for evaluative purposes, a change score greater than the estimated MDC is required as evidence of true change.[21] The MDC_{95} of 12.2s for the 10mFWT was large given the mean time of 11.4s to complete the distance, and reflects directly the issues with the test's measurement accuracy.

In contrast, 6MWT MDCs observed in this study (i.e. from 47–64m) are substantially lower than the MDC of 97.8m (ICC 0.87, SEM 35.3) from a reliability study with a sample of 550 healthy adults,[16] and the MDC of 125.6m (ICC 0.88, SEM 45.3) with a sample of adults with chronic lung disease.[17] This difference in MDC might also have been partially due to the slower walking speeds of children in the present study compared with adults. The observed differences in 6MWT MDCs among GMFCS levels suggest that a level-specific MDC should be used when interpreting change scores.

From a clinical interpretation standpoint, the observed MDC and LOA estimates require a 6MWT distance change that may be beyond that associated with therapeutic interventions in CP. Since research reporting 6MWT distance changes in children with CP could not be found, studies involving changes in walking speed were used to determine the applicability of this MDC. For example, one could speculate that this MDC is too large to detect change score differences of 0.11m/s (i.e. a 39.6m increase in 6min assuming a consistent pace over the 6-min distance) observed in a study comparing barefoot versus shoes/orthosis conditions in children with CP.[26] However, in a gait evaluation following BoNT-A injections of gastrocnemius muscles in children in GMFCS Level I, there were mean velocity gains of 0.23m/s.[27] This would translate into an increase of 82m on the 6MWT. It would also permit detection of the postrhizotomy gains of 0.26m/s observed by Wright et al.[28] in children with spastic diplegia.

Conclusion

In summary, the findings show that there are concerns with the test–retest reliability of the 10mFWT using the chosen testing protocol. In contrast, the 6MWT demonstrated excellent test–retest reliability with narrow 95%CIs. While adaptation of the standardized 6MWT instructions and encouragement was necessary for younger participants, the comparison of test–retest reliability between participants with testing issues and those without revealed that reliability was maintained despite these issues.

Further study with a larger sample is suggested to confirm the impact of GMFCS level and age on reliability. Using the power contours devised by Donner and Elaisziw[29] for reliability study sample size calculations, one would need at least 40 participants within a single GMFCS level or age stratum to be confident of reliability of 0.90 (where H_o=0.80) at an 80% power level. The MDC may not represent a minimal clinically important difference (MCID) in relation to a particular intervention or client group.[21] Formal work is required to establish MCIDs[17,28] for the 6MWT in children with CP. This could be done by means of a data-driven approach based on SEMs and observed change scores associated with common interventions in conjunction with consensus methods that would provide a clinically-driven impression of important change.[30]

Accepted for publication 6th November 2007.

Acknowledgements

This research was completed by the first five authors in partial fulfilment of the requirements of an MScPT degree at the University of Toronto.

All of the authors would like to acknowledge Linda Patrick (BSc PT, Physiotherapist, Bloorview Kids Rehab) and Christina Hay (BSc, Research Assistant, Bloorview Research Institute) for their invaluable assistance with the testing sessions, and the Bloorview Research Institute for supporting their time. The authors also thank the children and families for the time that they gave to participate in the study. The senior author, Virginia Wright, is jointly supported by the Bloorview Research Institute and a Career Development Award (2007 to 2011) through the Canadian Child Health Clinician Scientist Program (CCHCSP), a Canadian Institutes of Health Research Strategic Training Initiative.

References

1. Steinwender G, Saraph V, Scheiber S, Zwick EB, Uitz C, Hackl K. Intrasubject repeatability of gait analysis data in normal and spastic children. *Clin Biomech (Bristol, Avon)* 2000; **15:** 134–39.
2. Wright VF, Liu G, Milne F. Evaluation of the reliability of measurement of time-distance parameters of gait: a comparison in children with juvenile rheumatoid arthritis and children with cerebral palsy. *Physiother Can* 1999; **51:** 191–200.
3. Watson MJ. Refining the ten-metre walking test for use with neurologically impaired people. *Physiotherapy* 2002; **88:** 386–97.
4. van Loo MA, Moseley AM, Bosman JM, de Bie RA, Hassett L. Test-re-test reliability of walking speed, step length and step width measurement after traumatic brain injury: a pilot study. *Brain Inj* 2004; **18:** 1041–48.
5. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002; **166:** 111–17.

6. van Hedel HJ, Wirz M, Dietz V. Assessing walking ability in subjects with spinal cord injury: validity and reliability of three walking tests. *Arch Phys Med Rehabil* 2005; **86:** 190–96.
7. Rossier P, Wade DT. Validity and reliability comparison of 4 mobility measures in patients presenting with neurologic impairment. *Arch Phys Med Rehabil* 2001; **82:** 9–13.
8. Holub C, Lamont M, Lombardo A, et al. The reliability of the 6-minute walk test in patients with end-stage renal disease. *Acute Care Perspect* 2002; **11:** 8–11.
9. Pankoff B, Overend T, Lucy D, White K. Validity and responsiveness of the 6-minute walk test for people with fibromyalgia. *J Rheumatol* 2000; **27:** 2666–670.
10. Mossberg KA. Reliability of a timed walk test in persons with acquired brain injury. *Am J Phys Med Rehabil* 2003; **82:** 385–90.
11. Fry DK, Pfalzer LA. Reliability of four functional tests and rating of perceived exertion in persons with multiple sclerosis. *Physiother Can* 2006; **58:** 212–19.
12. Li AM, Yin J, Yu CCW, et al. The six-minute walk test in healthy children: reliability and validity. *Eur Respir J* 2005; **25:** 1057–60.
13. Lelieveld OT, Takken T, van der Net J, van Weert E. Validity of the 6-minute walking test in juvenile idiopathic arthritis. *Arthritis Rheum* 2005; **53:** 304–7.
14. Paap E, van der Net J, Helders PJ, Takken T. Physiologic response of the 6-minute walk test in children with juvenile idiopathic arthritis. *Arthritis Rheum* 2005; **53:** 351–56.
15. Nixon PA, Joswiak ML, Fricker FJ. A 6-minute walk test for assessing exercise tolerance in severely ill children. *J Pediatr* 1996; **129:** 362–66.
16. Tsang RC. Reference values for 6-minute walk test and hand-grip strength in healthy Hong Kong Chinese adults. *Hong Kong Physiother* 2005; **23:** 6–12.
17. Wise RA, Brown CD. Minimal clinically important differences in the 6-minute walk test and the incremental shuttle walking test. *COPD* 2005; **2:** 125–29.
18. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol* 1997; **39:** 214–23.
19. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. Norwalk, CT: Appleton and Lange, 1993.
20. Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use. Second edition. Oxford: Oxford University Press, 2000.
21. Stratford PW. Getting more from the literature: estimating the standard error of measurement from reliability studies. *Physiother Can* 2004; **56:** 27–30.
22. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* 1998; **26:** 217–38.
23. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **1:** 307–10.
24. Evans MD, Goldie PA, Hill KD. Systematic and random error in repeated measurements of temporal and distance parameters of gait after stroke. *Arch Phys Med Rehabil* 1997; **78:** 725–29.
25. Verschuren O, Takken T, Ketelaar M, Gorter JW, Helders PJ. Reliability for running tests measuring for agility and anaerobic muscle power in children and adolescents with cerebral palsy. *Pediatr Phys Ther* 2007; **19:** 108–15.
26. Buckon CE, Thomas SS, Jakobson-Huston S, Moor M, Sussman M, Aiona M. Comparison of three ankle-foot orthosis configurations for children with spastic diplegia. *Dev Med Child Neurol* 2004; **46:** 590–98.
27. Papadonikolakis AS, Vekris MD, Korompilias AV, Kostas JP, Ristanis SE, Soucacos PN. Botulinum A toxin for treatment of lower limb spasticity in cerebral palsy. *Acta Orthop Scand* 2003; **74:** 749–55.
28. Wright FV, Sheil E, Drake JM, Wedge JH, Naumann S. Evaluation of selective dorsal rhizotomy for the reduction of spasticity in cerebral palsy: a randomized controlled trial. *Dev Med Child Neurol* 1998; **40:** 239–47.
29. Donner A, Eliasziw M. Sample size requirements for reliability studies. *Stats Med* 1987; **6:** 441–48.
30. Wells G, Beaton D, Shea B, et al. Minimal clinically important differences: review of methods. *J Rheumatol* 2001; **28:** 406–12.

List of abbreviations

10mFWT	10-metre fast walk test
6MWT	6-minute walk test
CV	Coefficient of variation
LOA	Limits of agreement
MDC	Minimum detectable change
TBI	Traumatic brain injury