RELIABILITY OF GAIT PERFORMANCE TESTS IN INDIVIDUALS WITH LATE EFFECTS OF POLIO

Ulla-Britt Flansbjer, PhD and Jan Lexell MD, PhD

Department of Rehabilitation Medicine, Skåne University Hospital, Lund, Orupssjukhuset, 221 85 Lund, Sweden (Flansbjer, Lexell), Division of Rehabilitation Medicine Department of Clinical sciences, Lund University, Lund, Sweden (Flansbjer, Lexell) and Department of Health Sciences, Luleå University of Technology, Luleå, Sweden (Lexell)

Short title: Gait performance tests in post-polio

Corresponding address: Ulla-Britt Flansbjer, PhD, Department of Rehabilitation medicine, Lund University Hospital, Orupssjukhuset, 221 85 Lund, Sweden. Phone: (+46) 413-55 66 35. Fax: (+46) 413-55 67 09. Email: ulla-britt.flansbjer@skane.se

The study was completed within the context of the Centre for Ageing and Supportive Environments (CASE) at Lund University, funded by the Swedish Research Council on Social Science and Working Life, and had received financial support from the Swedish Association of Survivors of Traffic Accidents and Polio (RTP), Stiftelsen för bistånd åt rörelsehindrade i Skåne, Alfred Österlunds stiftelse and Skane county council's research and development foundation.

ABSTRACT

Objective: To assess the reliability of 4 gait performance tests in individuals with late effects of polio.

Design: An intra-rater (between occasions) test-retest reliability study.

Settings: University hospital.

Participants: Thirty men and women (mean age 63 ±6.4 years) with clinically and electrophysiologically verified late effects of polio.

Intervention: Not applicable.

Main outcome measures: The Timed "Up & Go" test, the Comfortable and the Fast Gait Speed tests, and the 6-Minute Walk test were assessed 7 days apart. Reliability was evaluated with the intraclass correlation coefficient ($ICC_{2,1}$), the mean difference between the test sessions ($\overline{d}$) and the 95% confidence intervals for $\overline{d}$, the standard error of measurement (SEM and SEM%), the smallest real difference (SRD and SRD%) and the Bland & Altman graphs.

Results: Test-retest agreements were high ($ICC_{2,1}$ 0.82-0.97) and measurement errors generally small. The standard error of measurement (SEM%), representing the smallest change that indicates a real (clinical) improvement for a group of individuals, was small (4% to 7%). The smallest real difference (SRD%), representing the smallest change that indicates a real (clinical) improvement for a single individual, was also small (12% to 21%).

Conclusion: These commonly used gait performance tests are highly reliable and can be recommended to evaluate improvements in various aspects of gait performance in groups of individuals as well as single individuals with late effects of polio.

INTRODUCTION

After an acute paralytic poliomyelitis infection, many individuals fully recovered or were left with only mild residual symptoms that did not affect their everyday life [1]. Several years later, after a stable period of at least 15 years, new symptoms, for example new weakness, muscle pain or atrophy, are experienced [2-5]. These symptoms are referred to as 'late effects of polio' or 'post-polio syndrome' (PPS). It is estimated that up to 80%, or 5 to 15 million people around the world, will need health care and rehabilitation over the next five decades as a result of their acute poliomyelitis infection. This makes it one of the most common neuromuscular conditions and a challenge to rehabilitation professionals [4]. As the late effects of polio often affect the lower limbs, many individuals experience increasing difficulties with their gait and mobility, which in turn affect their performance of everyday activities and restrict their participation [6-11]. Interventions aiming at improving gait performance is therefore one of the most common actions in the rehabilitation of these individuals [4].

To assess gait performance and evaluate changes following an intervention, we need safe and sound measurement tools [12]. In particular, measurement tools must be reliable, providing consistent measurements with low or no errors [13]. To comprehensively evaluate reliability, several statistical methods and indices are required that cover agreements between measurements, systematic changes in the mean and measurement errors [14]. Moreover, a measurement tool may be considered reliable but is not sensitive enough to detect a real (clinical) improvement following an intervention. A further advantage of a comprehensive reliability analysis is that it can be used to define limits for the smallest change that indicate an improvement, both for a group of patients and for individual patients.

Different gait performance tests have been used in patients with neurological conditions [15] and in individuals with late effects of polio [10, 16-26]. Some of these tests have been analysed for intra-rater reliability [18, 19, 23, 24], but even though they were found to be reliable, the statistical analyses were not sufficiently comprehensive and only one or a few of the tests were evaluated in each study. Further studies of the reliability of gait performance tests in individuals with late effects of polio are therefore needed.

The overall aim of this study was to assess the reliability of gait performance tests in individuals with late effects of polio. A set of statistical methods was used to evaluate the intra-rater (between occasions) test-retest reliability of 4 different gait performance tests and limits were defined for the smallest change that indicate a real (clinical) improvement, both for a group of individuals or a single individual.

MATERIAL AND METHODS

Thirty community-dwelling ambulant individuals (19 men and 11 women; mean age 63 years SD 6.4, 51-77 years; mean time since onset of new symptoms 14 years SD 5.5, 3-25 years; able to walk at least 300 m with or without an assistive device and/or an orthotics) were recruited from a post-polio rehabilitation clinic in a university hospital. All individuals met the criteria of postpoliomyelitis syndrome, as defined by Halstead and Rossi [27]. They all had a confirmed history of acute poliomyelitis affecting the lower limbs, new symptoms after a period of functional stability and no other diseases that could explain their reduced muscle strength. An electromyogram (EMG) had been recorded in the lower limbs as part of the initial routine clinical examination and verification of prior polio. On conventional EMG, motor unit action potentials in limbs previously affected with paralytic polio are abnormally enlarged and polyphasic in configuration and there is a decreased recruitment secondary to a reduction in the number of motor units available for activation during voluntary muscle contraction [4]. Each lower limb was classified according to the National Rehabilitation Hospital (NRH) Post-Polio Limb Classification [2]. All individuals had post-polio class III-V (indicating a clinically stable polio, clinically unstable polio or severely atrophic polio) in at least one of the lower limbs and 18 subjects in both lower limbs. Prior to inclusion, information about the purpose of the study was provided and each individual gave their written informed consent to participate.

Gait performance tests

Each individual underwent the following 4 gait performance tests: the Timed “Up & Go” test (TUG), the Comfortable and the Fast Gait Speed tests (CGS and FGS), and the 6-Minute Walk test (6MW). These tests are commonly used in patients with neurological conditions [15] and have also been used in our research group to evaluate gait performance in stroke patients before and after an intervention [28, 29].

The TUG [30] was developed primarily to evaluate basic functional mobility in frail elderly persons and has been used in PPS [18, 25, 26]. For the TUG, the individuals sat in a chair placed at the end of a marked 3-m walkway. They were instructed to sit with their back against the chair, and on the word “go”, stand up, walk at a comfortable speed past the 3-m mark, turn around, walk back and sit down in the chair. Each individual did one trial to become familiar with the test, and then performed the TUG twice with a one-min rest between each trial. The time from the start until the individual sat down in the chair with back support was measured and the mean of the two tests was recorded.

Gait speed timed over short distances (mostly 5 to 10 m) has been used as a determinant of mobility in PPS [10, 21, 23]. For the CGS and FGS a 14 m walkway was marked on the floor and the individuals were timed over the middle 10 m. For the CGS, the individuals were told to walk at a self-selected comfortable pace. For the FGS, the individuals were told to walk as fast and safely as possible without running. They started with the CGS three times in succession and with 30 s between each trial. After a further 30 s rest they continued with the FGS, also three times in succession, with 30 s between each trial. The time (in seconds) taken to walk 10 m was recorded for each trial and the mean value for CGS respectively FGS was calculated.

The 6MW is commonly used to assess patients with cardiovascular or cardiorespiratory problems [31] and is regarded as a submaximal test of aerobic capacity. The 6MW has also been used in PPS [17, 24]. For the 6MW, the individuals were instructed to walk 30 m between two marks on the floor. After passing either mark, they were told to turn and walk back. They were instructed to cover as much ground as possible and to walk as far as possible during six minutes, and were allowed to rest and then to continue walking; no subject had to rest during the test. They were informed when three minutes of the test remained. The 6MW was done once and the numbers of 30 m-lengths were counted. One wall was marked every meter so that the distance walked could be measured to the nearest meter.

Procedure

The individuals were tested on 2 occasions at the same time of the day and 7 days apart. All individuals were provided transport free of charge to and from the test site. The tests were performed in the following order: the TUG, the CGS, the FGS, and the 6MW. Individuals rested on a chair for 5 min before and after the 6MW. The same physiotherapist (UBF) supervised all tests and no verbal encouragement was given during the tests. The total time for the gait performance tests was approximately one hour. Throughout each session, individuals wore comfortable shoes and were allowed to use their assistive and/or orthotics. Seven individuals wore a drop-foot orthotics and one a walking aid (a stick). A digital stopwatch with an accuracy of one decimal figure in units of 1 sec was used to measure time. All tests were done in a 2.2 m wide corridor with a linoleum floor in a quiet part of the hospital.

Data and statistical analysis

All recruited individuals completed the tests and the statistical analyses are therefore based on N=30. The 4 recorded variables from the gait performance tests, obtained from the two test sessions, were used in the analysis. Descriptive statistics (means ± SD) were calculated for the characteristics of the 30 individuals. The difference between men and women for each variable was assessed with the two-sided *t*-test. The relationship between the 4 variables in each of the 2 test sessions was calculated using Pearson's correlation coefficient. There were no significant differences between the sexes for the clinical characteristics or the gait performance tests, so the results from the men and the women are combined throughout.

To determine the test-retest reliability several statistical methods were applied; a detailed account of these analyses, a rationale for their use and all equations have been presented previously [14]. Agreement between measurements was assessed with the intraclass correlation coefficients ($ICC_{2,1}$) and the mean difference between the test sessions ($\bar{d}$) together with the 95% confidence intervals for $\bar{d}$ (95% CI). Measurement errors were assessed by the standard error of measurement, SEM, and the SEM%. The smallest real difference, SRD, was calculated together with an 'error band' around the mean difference of the two measurements, $\bar{d}$, and the SRD%. The Bland-Altman graphs were formed to give a visual interpretation of the data as well as to determine reproducibility bias. A significance level smaller than 0.05 represented statistical significance. All calculations were performed using the SPSS 16.0 Software for Windows (SPSS Inc., Chicago, Ill., USA).

RESULTS

The means, standard deviations and the ranges of values for the 4 gait performance tests from the two test sessions are presented in Table 1. The differences between the means of the two tests were less than 6%. The mean values were relatively high compared to age-matched healthy individuals varying from 64% (CGS) to 86% (6MW) [32].

[Insert Table 1 about here]

The correlation between the 4 gait performance tests are presented in Table 2. There was a significant relationship between all 4 tests within each test session. The absolute values of Pearson's correlation coefficient (r) were in the range 0.53 to 0.88 ($p<0.01$).

[Insert Table 2 about here]

The reliability of the 4 gait performance tests are summarised in Table 3. For all tests, improvements are throughout presented as positive values. The $ICC_{2,1}$ values ranged from 0.82 to 0.97 and the 95% confidence intervals for $ICC_{2,1}$ from 0.65 to 0.98. All $\bar{d}$ values were close to zero and the widths of the 95% CI for $\bar{d}$ were narrow. For 3 of the 4 tests a positive value of $\bar{d}$ indicated that the performance at the second test session was better than at the first. In 2 of the 4 tests (TUG and 6MW), zero was not included in the 95% CI of $\bar{d}$, indicating that the better performance in the second test session was significant ($p<0.05$).

Measurement errors are presented both as the SEM (absolute values) and as the SEM% (independent of the units of measurement). The SEM%, indicating the measurement errors for a group of individuals, was low ranging from 4% to 7%. The 95% SRD indicates the range of measurement errors for a single individual. The SRD%, which represents the difference in relative terms, ranged from 12% to 21%.

[Insert Table 3 about here]

In Figure 1 the systematic variation around the zero line for all tests was revealed by forming ‘Bland & Altman graphs’. There were generally more values above the zero line for TUG and 6MW, illustrating the better performance at the second test session, but there were no other discernible pattern detected.

[Insert Figure 1 about here]

DISCUSSION

Individuals with late effects of polio often experience new muscle weakness, muscle atrophy, pain and fatigue in the lower limbs. This affects their mobility and gait performance [6-8, 21] and can impact on the performance of everyday activities and restrict their participation. To follow deterioration in gait performance longitudinally or changes before and after an intervention, we need robust measurement tools. In this study, the reliability of 4 commonly used gait performance tests was evaluated. We found that these tests were highly reliable and that relatively small differences are sufficient to detect real changes over time or improvements after an intervention.

Over the last decade, the assessment of reliability has developed from using only correlation coefficients to more comprehensive sets of statistical methods. It is now generally agreed that several statistical methods and indices are required to fully evaluate reliability [14]. Moreover, the concept of reliability has been expanded and the data from the analyses can be used to define limits for the smallest change that indicate a real improvement, both for a group of individuals and for a single individual [14, 33].

Several factors can influence reliability – in the present study, two factors were identified: the sample size and the test protocol. A sample size of 30 is considered sufficient to assess the reliability of a measurement and was therefore used here [34, 35]. It is well known that errors in the test protocol will affect reliability. Great care was taken to standardise the tests and to carefully follow the protocol, having the same time interval between the tests and using the same commands, to optimize the reliability. Thus, with all conditions as stable as possible, any variability between the two test sessions is taken to represent the variability in the measurement parameters within and between the individuals.

Several tests have been used to evaluate gait performance in individuals with late effects of polio [10, 16-26]. The 4 tests evaluated in the present study were selected as they are commonly used clinically and very easy to administer [15]. Also, they cover a variety of aspects, such as velocity and endurance, which provide a comprehensive picture of the complexity of gait performance. Not surprisingly, there was a significant correlation between the different tests, as they measure similar, although in some ways different, aspects of gait performance. The high correlation between CGS and FGS indicates that gate speed measured in different ways is highly related. Similarly, the negative correlations between 6MW and CGS, FGS and TUG indicate that gait velocity was related to gait endurance in these individuals; the faster a person could perform during the FGS, CGS or TUG the longer he or she could walk during the 6MW.

The ICC is by far the most common method to evaluate reliability and it has also been used to assess the reliability of gait performance tests in individuals with late effects of polio. These studies have, in agreement with the present study, found fairly high ICC values. One study of the 6MW (N=19) [24] reported ICC values from 0.90 to 0.98. In a study of self preferred and maximal gait speed (N=63) [19] the ICC-values varied from 0.94 to 0.97. In the study by Brehm et al. [23] (N=14), walking during 4 to 5 minutes (with a lightweight portable gas analyse system to determine the energy demands), was evaluated and the ICC values were 0.93 to 0.99. Finch et al. [18] (N=15) tested walking on a treadmill and values were reported to be 0.85. These result are also in agreement with the test-retest reliability in patients with other neurological conditions (ICC 0.93 to 0.97) [36] and after stroke (ICC 0.94 to 0.99) [28].

Using only the ICC can lead to incorrect conclusions about the reliability of a measurement [14]. ICC assesses the agreement between repeated measurements and thereby only the variance between individuals. Any variability in the measurements within individuals must be considered. This is clearly visualized by using the 'Bland & Altman analyses' and the 'Bland & Altman graphs'. For example, the better performances during the second test session can be seen. This suggests a learning effect for 3 of the 4 tests and a significant effect for TUG and 6MW; for 6MW there are similar findings in healthy individuals and after stroke [28, 37]. However, the mean differences between the two sessions ($\bar{d}$) were close to zero and the confidence intervals were narrow (cf. Table 3). Although, this indicates that the learning effect was small, it may have to be taken into account in future studies of this population.

Two indices, SEM and the SEM%, were used here to evaluate the measurement errors. The SEM gives the measurement errors in absolute values, whereas the SEM% is independent of the units of measurement, and therefore more easily interpreted. The SEM% represents the limit for the smallest change that indicates a real improvement for a group of individuals over time or following an intervention. All SEM% values in the present study were below 8%, which implies that the tests are sensitive and can be used to detect small, clinically relevant, changes in gait performance in this population.

The data can also be used to determine whether a method is sufficiently sensitive to detect a real change for a single individual. The smallest real difference (SRD), [14, 33], is as a method linking reproducibility to responsiveness. The SRD% is independent of the unit of measurement and, like the SEM%, more easily interpreted. For the gait performance tests in this study, the size of the relative change (SRD%) should exceed 12% (6MW) up to 21% (TUG) to indicate a real change. From the data in Table 3, the relative improvement (SRD%) needed to detect such a change can be calculated for any subject in our study. For example, on average an individual in the present study covered 478 during the 6MW, and had to walk a further 57 m to indicate a real change. This is more or less the same as in post-stroke patients, where an improvement of 51 m during the 6MW indicated a real change [28]. Only 2 studies have evaluated measurement errors for gait performance tests in individuals with late effects of polio. Brehm et al. [23] reported better reliability with increased numbers of repetitions (from 1 to 4) reducing the SEM% from 4.5 to 2.2 and the SRD% from 13% to 6.1%. Horeman et al.[19] assessed the reliability of self-preferred and maximal walking speed and determined the measurement errors using the Limits of Agreement (LOA), algebraically similar to SRD [38]. These authors found a percentage change from mean of about 15%, indicating the smallest change within an individual.

Limitations

The individuals assessed were all functioning well, with an ability to walk at least 300 m and with a predicted gait performance close to healthy age-matched individuals [32]. The results in this study should only be generalized to those individuals with late effects of polio that have a fairly maintained walking ability.

CONCLUSIONS

All 4 gait performance tests in this study showed: i) high agreement between the test-retest measurements; ii) no substantial systematic changes in the mean and small measurement errors; and iii) sufficient sensitivity to enable the detection of real (clinical) changes in measurement score. Taken together, these tests can be recommended in clinical practice as well as research to evaluate various aspects of gait performance and changes over time in individuals with late effects of polio. As all the tests are highly related, the choice of test depends on what aspect of gait performance that is being evaluated.

REFERENCES

1. Agre JC, Rodriquez AA, Sperling KB: Symptoms and clinical impressions of patients seen in a postpolio clinic. Arch Phys Med Rehabil 1989;70:367-370.

2. Halstead LS, Gawne AC, Pham BT: National rehabilitation hospital limb classification for exercise, research, and clinical trials in post-polio patients. Ann N Y Acad Sci 1995;753:343-353.

3. Farbu E, Rekand T, Gilhus NE: Post-polio syndrome and total health status in a prospective hospital study. Eur J Neurol 2003;10:407-413.

4. Trojan DA, Cashman NR: Post-poliomyelitis syndrome. Muscle Nerve 2005;31:6-19.

5. Farbu E, Gilhus NE, Barnes MP, et al.: EFNS guideline on diagnosis and management of post-polio syndrome. Report of an EFNS task force. Eur J Neurol 2006;13:795-801.

6. Grimby G, Jonsson AL: Disability in poliomyelitis sequelae. Phys Ther 1994;74:415-424.

7. Ivanyi B, Nollet F, Redekop WK, et al.: Late onset polio sequelae: disabilities and handicaps in a population-based cohort of the 1956 poliomyelitis outbreak in The Netherlands. Arch Phys Med Rehabil 1999;80:687-690.

8. Nollet F, Beelen A, Prins MH, et al.: Disability and functional assessment in former polio patients with and without postpolio syndrome. Arch Phys Med Rehabil 1999;80:136-143.

9. Thoren-Jonsson AL: Adaptation and ability in daily occupations in people with poliomyelitis sequelae. University of Gothenburg; 2000.

10. Willen C, Thoren-Jonsson AL, Grimby G, Sunnerhagen KS: Disability in a 4-year follow-up study of people with post-polio syndrome. J Rehabil Med 2007;39:175-180.

11. Larsson Lund M, Lexell J: Perceived participation in life situations in persons with late effects of polio. J Rehabil Med 2008;40:659-664.

12. Holden MK, Gill KM, Magliozzi MR, Nathan J, Piehl-Baker L: Clinical gait assessment in the neurologically impaired. Reliability and meaningfulness. Phys Ther 1984;64:35-40.

13. Rothstein JM: Measurement and clinical practice: Theory and application. In: Rothstein JM, ed. Measurement in Physical Therapy. New York: Churchill Livingstone, 1985; 1-46.

14. Lexell JE, Downham DY: How to Assess the Reliability of Measurements in Rehabilitation. Am J Phys Med Rehabil 2005;84:719-723.

15. Tyson S, Connell L: The psychometric properties and clinical utility of measures of walking and mobility in neurological conditions: a systematic review. Clin Rehabil 2009;23:1018-1033.

16. Perry J, Mulroy SJ, Renwick SE: The relationship of lower extremity strength and gait parameters in patients with post-polio syndrome. Arch Phys Med Rehabil 1993;74:165-169.

17. Noonan VK, Dean E, Dallimore M: The relationship between self-reports and objective measures of disability in patients with late sequelae of poliomyelitis: a validation study. Arch Phys Med Rehabil 2000;81:1422-1427.

18. Finch LE, Venturini A, Mayo NE, Trojan DA: Effort-limited treadmill walk test: reliability and validity in subjects with postpolio syndrome. Am J Phys Med Rehabil 2004;83:613-623.

19. Horemans HL, Beelen A, Nollet F, Lankhorst GJ: Reproducibility of walking at self-preferred and maximal speed in patients with postpoliomyelitis syndrome. Arch Phys Med Rehabil 2004;85:1929-1932.

20. Willen C, Sunnerhagen KS, Ekman C, Grimby G: How is walking speed related to muscle strength? A study of healthy persons and persons with late effects of polio. Arch Phys Med Rehabil 2004;85:1923-1928.

21. Horemans HL, Bussmann JB, Beelen A, Stam HJ, Nollet F: Walking in postpoliomyelitis syndrome: the relationships between time-scored tests, walking in daily life and perceived mobility problems. J Rehabil Med 2005;37:142-146.

22. Sorenson EJ, Daube JR, Windebank AJ: A 15-year follow-up of neuromuscular function in patients with prior poliomyelitis. Neurology 2005;64:1070-1072.

23. Brehm MA, Nollet F, Harlaar J: Energy demands of walking in persons with postpoliomyelitis syndrome: relationship with muscle strength and reproducibility. Arch Phys Med Rehabil 2006;87:136-140.

24. Gylfadottir S, Dallimore M, Dean E: The relation between walking capacity and clinical correlates in survivors of chronic spinal poliomyelitis. Arch Phys Med Rehabil 2006;87:944-952.

25. Gonzalez H, Sunnerhagen KS, Sjoberg I, Kaponides G, Olsson T, Borg K: Intravenous immunoglobulin for post-polio syndrome: a randomised controlled trial. Lancet Neurol 2006;5:493-500.

26. Lehmann K, Sunnerhagen KS, Willen C: Postural control in persons with late effects of polio. Acta Neurol Scand 2006;113:55-61.

27. Halstead LS, Rossi CD: New problems in old polio patients: results of a survey of 539 polio survivors. Orthopedics 1985;8:845-850.

28. Flansbjer U-B, Holmbäck AM, Downham D, Patten C, Lexell J: Reliability of gait performance tests in men and women with hemiparesis after stroke. J Rehabil Med 2005;37:75-82.

29. Flansbjer UB, Miller M, Downham D, Lexell J: Progressive resistance training after stroke: Effects on muscle strength, muscle tone, gait performance and perceived participation. J Rehabil Med 2008;40:42-48.

30. Podsiadlo D, Richardson S: The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39:142-148.

31. Bittner V, Weiner DH, Yusuf S, et al.: Prediction of mortality and morbidity with a 6-minute walk test in patients with left ventricular dysfunction. SOLVD Investigators. Jama 1993;270:1702-1707.

32. Steffen TM, Hacker TA, Mollinger L: Age- and gender-related test performance in community-dwelling elderly people: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds. Phys Ther 2002;82:128-137.

33. Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL: Smallest real difference, a link between reproducibility and responsiveness. Qual Life Res 2001;10:571-578.

34. Hopkins WG: Measures of reliability in sports medicine and science. Sports Med 2000;30:1-15.

35. Donner A, Eliasziw M: Sample size requirements for reliability studies. Stat Med 1987;6:441-448.

36. 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.

37. Camarri B, Eastwood PR, Cecins NM, Thompson PJ, Jenkins S: Six minute walk distance in healthy subjects aged 55-75 years. Respir Med 2006;100:658-665.

38. Bland JM, Altman DG: Measuring agreement in method comparison studies. Stat Methods Med Res 1999;8:135-160.

LEGENDS

Figure 1. The differences between test sessions 2 and 1 (test 2 minus test 1) plotted against the means of the two test sessions for the 4 gait performance tests. From these ‘Bland & Altman graphs’, the systematic variation around the zero line was revealed.

Table 1. *Results from the 4 gait performance tests for the 30 men and women with late effects of polio*

	Test session 1	Test session 2
Timed Up & Go (sec)	10.7 (1.8), 7.1-14.7	10.1 (1.6), 6.1-13.1
Gait speed (sec)		
Comfortable Gait Speed	9.0 (1.6), 6.9-13.5	8.7 (1.3), 6.4-11.5
Fast Gait Speed	6.2 (1.3), 3.4-9.9	6.2 (1.3), 3.6-9.0
6-Minutes Walk (m)	469 (89), 334-699	486 (91), 347-694

Values are presented as the mean (SD) and range, obtained from each test session.

Table 2. *Pearson's correlation coefficients between the gait performance tests for the 30 men and women with late effects of polio*

		Test session 2			
		TUG	CGS	FGS	6MW
Test session 1	TUG	—	.61	.58	-.54
	CGS	.59	—	.78	-.80
	FGS	.60	.69	—	-.88
	6MW	-.53	-.71	-.83	—

Abbreviations: Timed Up & Go (TUG), Comfortable Gait Speed (CGS), Fast Gait Speed (FGS), 6-Minutes Walk test (6MW). All correlation coefficients were significant ($p<0.01$).

Table 3. *Reliability of the 4 gait performance tests for the 30 men and women with late effects of polio*

Test	$ICC_{2.1}$	95% CI for ICC	$\bar{d}$	95% CI for $\bar{d}$	SEM	SEM%	95% SRD	SRD%
Timed Up & Go (seconds)	0.85	0.72 to 0.93	0.61	0.27 to 0.95	0.77	7.4	-1.53 to 2.75	20.6
Gait speed (seconds)								
Comfortable Gait Speed	0.82	0.65 to 0.91	0.24	-0.09 to 0.57	0.64	7.2	-1.53 to 2.01	20.0
Fast Gait Speed	0.93	0.87 to 0.97	-0.01	-0.17 to 0.15	0.32	5.2	-0.91 to 0.89	14.5
6-Minutes Walk (metres)	0.97	0.93 to 0.98	17	8 to 25	19	4.2	-38 to 72	11.6

Abbreviations: $ICC_{2.1}$ = intraclass correlation coefficient; CI = confidence interval; SEM = standard error of measurement; SRD = smallest real difference.

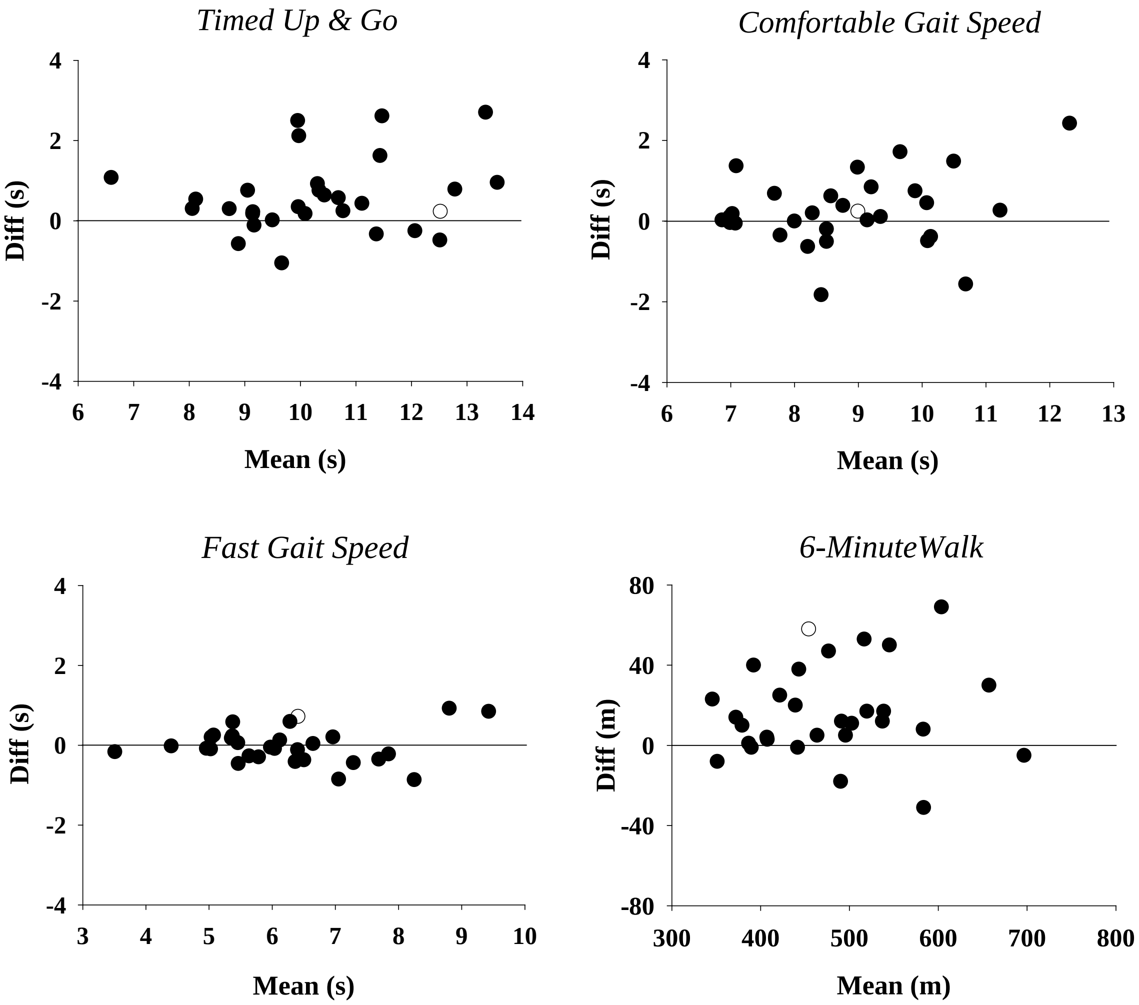
Timed Up & Go
Diff (s)
Mean (s)
Comfortable Gait Speed
Diff (s)
Mean (s)
Fast Gait Speed
Diff (s)
Mean (s)
6-MinuteWalk
Diff (m)
Mean (m)

Figure 1