

The reliability and validity of the L-test in people with Parkinson's disease

Bernhard Haas*, Elizabeth Clarke, Laura Elver, Emma Gowman,
Emma Mortimer, Erin Byrd

University of Plymouth, School of Health Professions, Peninsula Allied Health Centre, Derriford Rd, Plymouth PL6 8BH, United Kingdom

Abstract

Objective To evaluate the test–retest and concurrent validity of the L-test in a group of participants with mild to moderate Parkinson's disease. The L-test is an extended version of the Timed up and Go test, incorporating a longer walking distance and turns in two directions.

Design Cross-sectional.

Setting Community.

Participants 16 participants (13 male), mean age 75 (SD 6.7) mean duration since diagnosis 7.1 years (± 2.8). Disease severity was mild to moderate on the Hoehn and Yahr scale (mean 2.1; mode 2; range 1–3). 14 participants (12 male) completed the study.

Interventions Not applicable.

Main outcome measures A Bland and Altman plot examined the agreement between first and second testing occasion of the L-test. Intra-class correlation coefficients (ICC) assessed the test–retest reliability. Concurrent validity was established by correlating the L-test with the Timed up and Go test (TUG). The Minimal Detectable Change with 95% confidence interval (MDC95) was calculated to determine the true change not due to chance.

Results The L-test showed excellent test–retest reliability on the Bland & Altman plot and the ICC. There was a high degree of agreement between measurements taken on days 1 and 2. The L-test correlated strongly with the Timed up and Go test on both measurement days with $r = 0.97$ ($p < 0.001$) and $r = 0.96$ ($p < 0.001$). The MDC95 was 5.31 seconds.

Conclusions The L-test is a reliable and valid outcome measurement for the assessment of walking ability in participants with mild to moderate Parkinson's disease.

© 2017 Chartered Society of Physiotherapy. Published by Elsevier Ltd. All rights reserved.

Keywords: Gait; Reliability; Parkinson's disease; Validity; Minimum Detectable Change

Introduction

Parkinson's disease (PD) is a progressive degenerative neurological conditions which leads to significant motor and non-motor symptoms [1,2]. Two of the major signs of PD, hypokinesia and rigidity in particular can have a negative effect on gait, balance and mobility due to the decline in rhythmic, co-ordinated movements [3]. Impairments of gait relate closely to activity limitations in this patient group [4]. The assessment and monitoring of gait function depends on the availability of valid and reliable outcome measurements.

In particular, simple tools which are easy to use can yield the greatest benefit in research and clinical practice. A number of tools have been developed that include timed tests of walking speed or walking endurance. However, not all of these tests have been validated for use in people with PD. The 10-m walk test has shown good repeatability in people with PD [5], however, this test only assesses simple straight walking and therefore has limitations in people who have specific difficulties in turning. Similarly, the 6-minute walk test has shown excellent repeatability [5], but taking the recommended minimum of two consecutive assessments may be a problem for some people with PD with limited walking endurance. The timed up and go (TUG) test is frequently used in the assessment of gait in a number of conditions, including

* Corresponding author.

E-mail address: bernhard.haas@plymouth.ac.uk (B. Haas).

stroke [6], older fallers [7] and in people with multiple sclerosis [8]. The TUG has also been used in studies involving people with PD [9–11]. It is highly recommended as a tool to assess mobility and falls risk due to its simplicity and ease of completion [12]. Longer TUG times are generally associated with decreased mobility and increased number of falls, but the standard TUG may not be sophisticated enough to differentiate sufficiently between individuals at the various stages of the disease progression in PD [13]. Reliability of the TUG has been found to be ‘good’ to ‘excellent’ in people with PD [14–16], but the TUG only assesses turning in a preferred direction where unilateral impairments may be masked by the opportunity to turn in a preferred direction [9]. Therefore, despite good reliability and validity the TUG has limitations in PD which may currently only be overcome by adding complex and potentially prohibitively expensive instrumentation. The L test is a modification of the TUG which extends the walking distance from 6 to 20 m and also requires participants to make turns in both a clockwise and anticlockwise direction. It therefore may have the potential to overcome limited sensitivity and ceiling effects. The L test was initially devised for people following lower limb amputation [17] and has since been evaluated in people following stroke [18] and hospitalised older people [9]. In all three studies the L test has shown good reliability and also correlated well with the TUG. The aim of this study was to assess the L test in a group of participants with PD in terms of test/re-test reliability and concurrent validity with the TUG.

Methods

Participants

In total, 16 participants with PD were recruited through the local Parkinson’s UK group and the Parkinson’s UK Research Support network (13 males). Their mean age (Standard Deviation) was 75 years (± 6.7) and their mean duration from diagnosis with PD was 7.1 years (± 2.8). Disease severity was mild to moderate on the Hoehn and Yahr scale [19] (mean 2.1; mode 2; range 1–3). Exclusion criteria were any significant co-morbidities affecting gait and balance, inability to walk 20 m without assistance or inability to follow the study protocol. Two participants were unable to attend the second testing session and therefore 14 participants (12 male) completed the study. The study was approved by the University Ethics committee and all people participated with informed written consent in accordance with the Declaration of Helsinki.

Outcome measurements

Timed up and Go test

The TUG is a commonly used measure of mobility. The test comprises of a rise from a standard chair, a 3 m walk, 180° turn around a cone or line before returning to the chair to sit down. The recorded time beginning on the word “go”

and ending when the participant’s back touches the back of the chair. We used a chair without armrests for all tests.

L test of functional mobility

The L-test is an extended version of the TUG, designed to assess advanced functional ambulation with further details provided elsewhere [18]. The time is recorded as the participant stands up from the chair, walks 3 m to a cone, turns 90° to the right, continues walking 7 m to the next cone, turns left 180° around the cone, then returns to the chair along the same path. The recorded time beginning on the word “go” and ending when the participant’s back touches the back of the chair. Fig. 1 shows schematic diagrams for the L-test and the TUG.

Experimental procedure

Participants completed two timed trials of both outcome measurements, with a 1-minute rest between trials and a 2-minute rest between measurements. This was performed on two occasions, seven days apart, to establish test–retest reliability. In order to control medication effects, the trials were completed at approximately the same time of day. Participants completed the TUG and L-test in a randomised order, which was kept consistent between the two occasions. Participants were instructed to walk at their comfortable walking speed, with or without any of their usual walking aids. Prior to participation, the participants received a standardised demonstration and explanation of both outcome measurements. For participant safety, chairs were placed in close proximity to the testing area. In addition, one of the researchers walked in close proximity to the participant without providing any encouragement or trying to set the walking pace. The mean times of both trials were used for data analysis.

Statistical analysis

Data was analysed using the Statistical Package for the Social Sciences (SPSS) (IBM Corp. SPSS Statistics for Windows, Version 22.0. Armonk, NY, USA). Data was tested for normality using Shapiro–Wilk test and this indicated that the data was not normally distributed. Therefore, we applied log transformation prior to data analysis. Test–retest reliability was determined using intraclass correlation coefficients (ICC) with a confidence interval (CI) of 95%. Reliability was considered poor if the ICC < 0.40 , moderate to good if $0.41–0.74$ and excellent if > 0.75 [20,21]. In addition, we then used the ICC results to calculate the Minimal Detectable Change (MDC at 95% confidence interval—MDC₉₅) of both the TUG and the L test using the following formula [22]:

$$\text{MDC}_{95} = 1.96 \times \text{SEM} \times \sqrt{2};$$

The Standard Error of Measurement (SEM) was calculated in the following way:

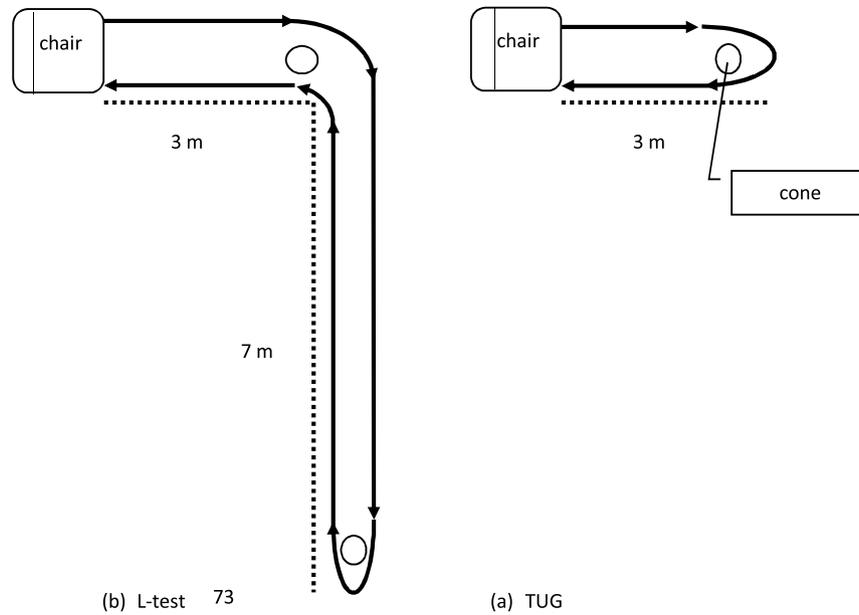


Fig. 1. Schematic diagram of the L-test (a) and the TUG (b).

$SEM = SD \times \sqrt{[1 - r]}$; (SD is the Standard Deviation of the baseline test and r is the ICC reliability coefficient).

Agreement of the L-test measurements between testing occasions one and two were also evaluated using the ‘limits of agreement’ method described by Bland and Altman [23].

Concurrent validity was examined by determining the correlation between the TUG and the L-test using Pearson’s correlation coefficients (r). Correlations were considered negligible if between 0 and 0.20, weak if 0.21–0.40, moderate if 0.41–0.60, strong if 0.61–0.80 and very strong if 0.81–1.00 [24]. A power calculation showed that 7 participants would be required to demonstrate a power $\beta = 0.97$ with a correlation of $r = 0.96$ and $p = 0.05$ [25].

Results

The mean (Standard Deviation—SD) time taken to complete the TUG on occasions one and two were 15.33 seconds (± 5.6) and 12.48 seconds (± 6.5) respectively. The mean (SD) time taken to complete the L-test on occasions one and two were 35.46 seconds (± 11.1) and 29.66 seconds (± 15.3) respectively.

Test–retest reliability and Minimally Detectable Change

Fig. 2 shows the Bland and Altman plot for the test–retest agreement between the L-test scores on day 1 and day 2. All data points are within two standard deviations and do not indicate a systematic bias. The mean difference between the two measurement days was 2.12 seconds. This also represents the bias or the difference between the mean difference between the two results and zero difference.

The ICC for the TUG was 0.92 (95% CI: 0.75 to 0.98) and the ICC for the L-test was 0.97 (95% CI: 0.86 to 0.99). Therefore, the test–retest reliability for both tests was determined to be excellent. The MDC_{95} for the TUG was 4.35 seconds and for the L-test was 5.31 seconds.

Concurrent validity

A positive significant correlation was found between the TUG and the L-test on both testing occasions. Pearson’s correlation coefficients were $r = 0.97$ ($p < 0.001$) and $r = 0.96$ ($p < 0.001$) for testing occasions one and two respectively. The correlations were therefore considered to be very strong. Figs. 3 and 4 show the relationships between the TUG and the L-test on testing occasions one and two respectively.

Discussion

Our study aimed to evaluate the reliability and concurrent validity of the L-test of functional mobility in a group of participants with PD. This is the first study to evaluate the L-test in this patient group. The ICC reliability coefficient for the TUG in our study was 0.95 (95% CI: 0.75 to 0.98) and this is regarded as ‘excellent’ according to published criteria [20,21]. Our ICC results for the TUG agree with the findings in other studies, which have also shown excellent reliability evaluations in participants with PD [5,14,26,27]. Our Bland and Altman plot of agreement suggests a high degree of agreement of the measurement scores between the two measurement occasions. The Bland and Altman plot does not show any systematic difference over the range of walking speeds. In addition, calculated the ICC from the results of two measurement occasions one week apart and at a similar time

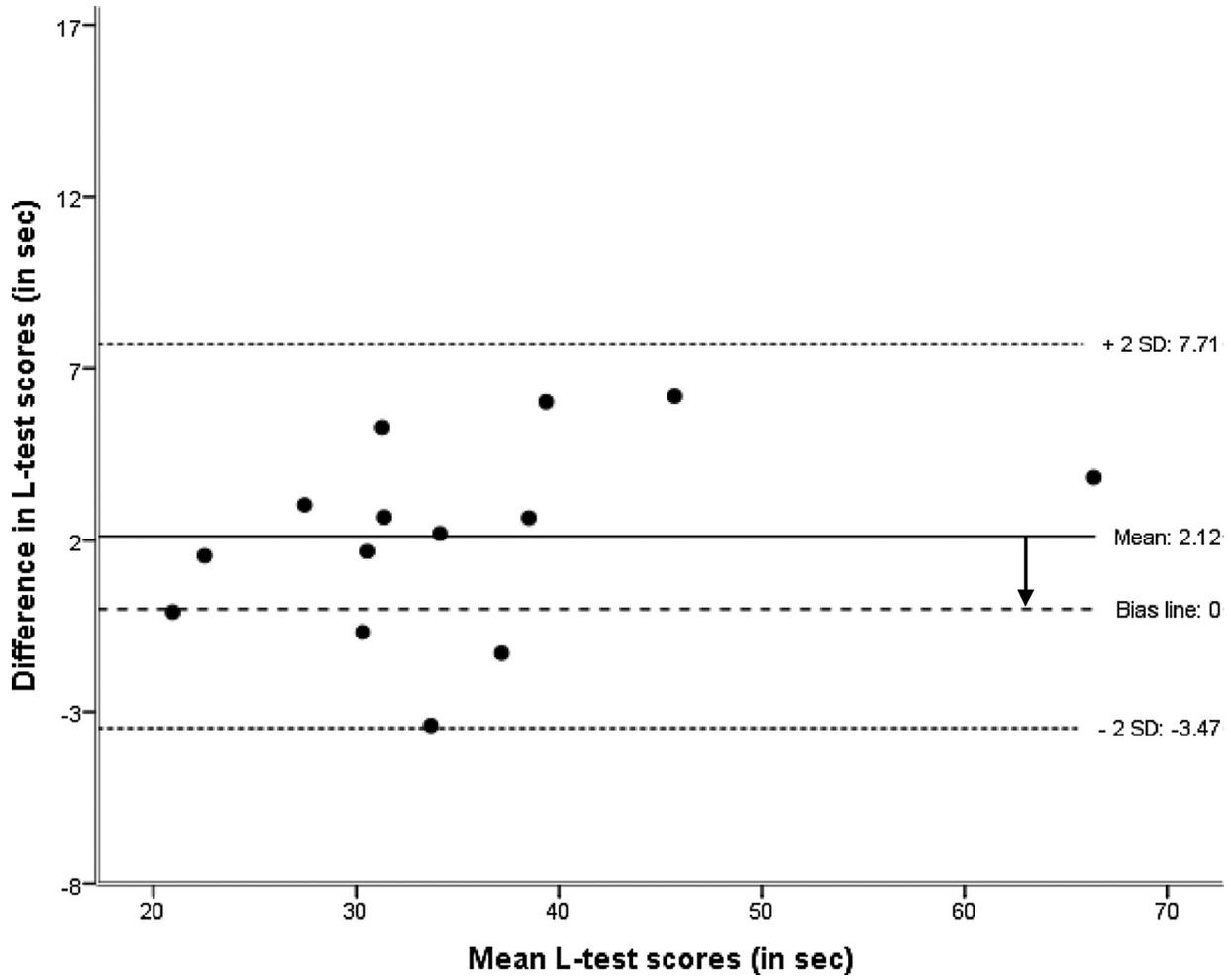


Fig. 2. Bland and Altman plot for test–retest agreement of L-test scores on day 1 and day 2.

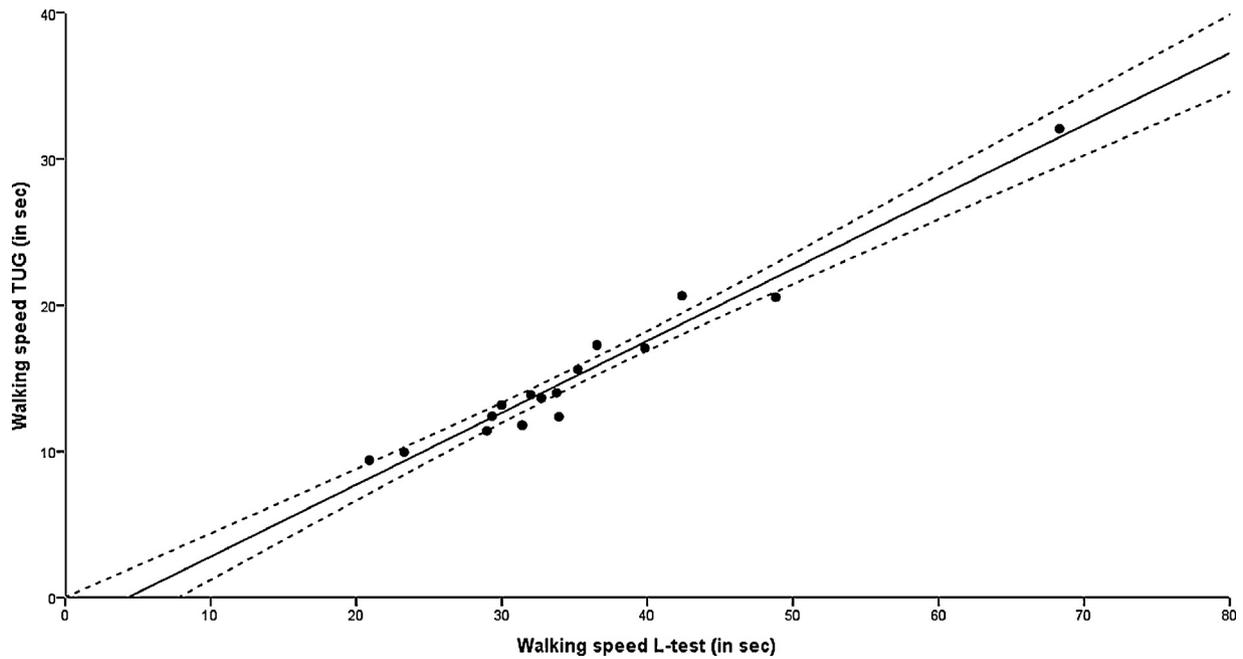


Fig. 3. Relationship between TUG and L-test on day 1.

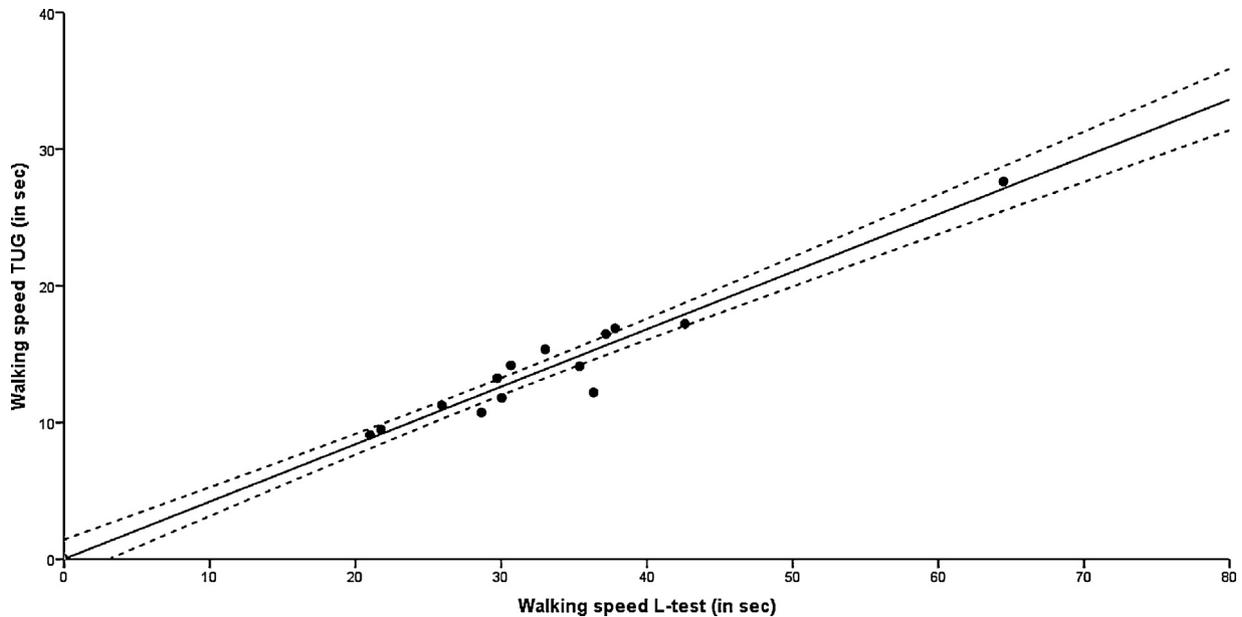


Fig. 4. Relationship between TUG and L-test on day 2.

of day. The results for the test–retest ICC for the L-test in our study was 0.97 (95% CI: 0.86 to 0.99) which also suggests excellent reliability. This suggests that the L-test should be reliable irrespective of walking speed [23]. We also calculated the ‘Minimal Detectable Change’ (MDC_{95}), and this was 5.31 seconds. The MDC is defined as the ‘minimal amount of change that is not likely to be due to chance’ [22]. For clinical practice, any improvements in walking speed following an intervention programme and measured by the L-test therefore need to be above 5.31 seconds, to be deemed clinically significant. We also calculated the MDC_{95} for the TUG, which was 4.35 seconds. We are not aware of any previous MDC_{95} calculations for the TUG in this patient population, so future studies using the TUG may need to consider that true improvements in walking speed may need to be above 4.35 seconds for this walking test.

Our study demonstrated high construct validity of the L-test; we used the well-established TUG test as our concurrent ‘gold standard’ measurement as it is well established in various patient populations and widely used in clinical practice [6–11]. Other studies have similarly evaluated the construct validity of the L-test by testing the correlation with the TUG; these studies have also shown high correlations in stroke survivors ($r = 0.89$) [18], participants with lower limb amputations ($r = 0.93$) [17] and in an elderly population in hospital ($r = 0.96$) [9]. Our correlation coefficients of $r = 0.96$ and $r = 0.97$ for testing occasions one and two respectively are therefore very much in line with these findings in other patient populations.

Our evaluation shows excellent reliability and validity of the L-test in this patient population. However, we are unable to demonstrate superiority or inferiority in comparison to the TUG. The need for a larger testing area may limited the use

to the L-test in some settings, but the added length and turns may be of interest in certain assessments.

Limitations

Our participants were mostly classified as Hoehn & Yahr stages 1 and 2 which corresponds to mild to moderate PD [19]. This is consistent with the level of disease progression that is observed within ambulatory, community dwelling people with PD [16]; however, we are unable to comment on the suitability of this test in participants with PD and more severe walking limitations. For safety reasons a researcher walked in close proximity of all participants in all trials. Whilst every effort was made that this should not influence the self-selected walking speed of the participant we cannot exclude this as a possibility. In addition, all trials were conducted in environments where potential distractions could not be excluded. Whilst this had the potential of influencing the results it may also be considered as relating more closely to a realistic measurement environment in a real clinical situation. We observed freezing of gait on a very small number of occasions and in both tests. However, we did not systematically record this, so therefore cannot comment which of these tests may trigger freezing more often.

Conclusion

We conclude that the L-test of functional mobility is a reliable and valid test in people with mild to moderate PD. The L-test provides an assessment of walking over a longer walking distance and involves turning in two directions, when compared to the TUG. For the assessment of changes in walk-

Key messages

- This study has established for the first time the reliability and validity of the L-test of functional mobility in participants with Parkinson's disease.
- The study also calculated the Minimum Detectable Change (MDC₉₅) of the L-test as well as the Timed up and Go test. The MDC₉₅ for the L-test was 5.31 seconds and 4.35 seconds for the Timed up and Go test.

ing ability a minimum change of 5.31 seconds is needed to demonstrate an actual change in performance in this patient group.

Conflicts of interest: None declared.

References

- [1] Cheon S-M, Ha M-S, Park MJ, Kim JW. Nonmotor symptoms of Parkinson's disease: prevalence and awareness of patients and families. *Parkinsonism Relat Disord* 2008;14(4):286–90.
- [2] Shulman LM, Taback RL, Bean J, Weiner WJ. Comorbidity of the nonmotor symptoms of Parkinson's disease. *Mov Disord* 2001;16(3):507–10.
- [3] Tolleson CM, Dobolyi DG, Roman OC, Kanoff K, Barton S, Wylie SA, et al. Dysrhythmia of timed movements in Parkinson's disease and freezing of gait. *Brain Res* 2015;1624:222–31.
- [4] Tan D, Danoudis M, McGinley J, Morris ME. Relationships between motor aspects of gait impairments and activity limitations in people with Parkinson's disease: a systematic review. *Parkinsonism Relat Disord* 2012;18(2):117–24.
- [5] Schenkman M, Cutson TM, Kuchibhatla M, Chandler J, Pieper C, Kim JS, et al. Reliability of impairment and physical performance measures for persons with Parkinson's disease. *Phys Ther* 1997;77(1):19–27.
- [6] van Bloemendaal M, van de Water A, van de Port I. Walking tests for stroke survivors: a systematic review of their measurement properties. *Disabil Rehabil* 2012;34(26):2207–21.
- [7] Schoene D, Wu SMS, Mikolaizak AS, Menant JC, Smith ST, Delbaere K, et al. Discriminative ability and predictive validity of the timed up and go test in identifying older people who fall: systematic review and meta-analysis. *J Am Geriatr Soc* 2013;61(2):202–8.
- [8] Sebastião E, Sandroff BM, Learmonth YC, Motl RW. Validity of the timed up and go as a measure of functional mobility in persons with multiple sclerosis. *Arch Phys Med Rehabil* 2016;97:1072–7.
- [9] Nguyen VC, Miller WC, Asano M, Wong RY. Measurement properties of the L test for gait in hospitalized elderly. *Am J Phys Med Rehabil* 2007;86(6):463–8.
- [10] Zampieri C, Salarian A, Carlson-Kuhta P, Aminian K, Nutt JG, Horak FB. The instrumented timed up and go test: potential outcome measure for disease modifying therapies in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 2010;81(2):171–6.
- [11] Kotagal V, Albin RL, Müller MLTM, Koeppel RA, Studenski S, Frey KA, et al. Advanced age, cardiovascular risk burden, and timed up and go test performance in Parkinson disease. *J Gerontol A Biol Sci Med Sci* 2014;69(12):1569–75.
- [12] American Geriatrics Society, British Geriatrics Society and American Academy of Orthopaedic Surgeons Panel on Falls Prevention. Guideline for the prevention of falls in older persons. *J Am Geriatr Soc* 2001;49:664–72.
- [13] Weiss A, Herman T, Plotnik M, Brozgol M, Maidan I, Giladi N, et al. Can an accelerometer enhance the utility of the timed up & go test when evaluating patients with Parkinson's disease? *Med Eng Phys* 2010;32(2):119–25.
- [14] Morris S, Morris ME, Ianssek R. Reliability of measurements obtained with the timed up & go test in people with Parkinson's disease. *Phys Ther* 2001;81(2):810–8.
- [15] Gusi N, Carmelo Adsuar J, Corzo H, del Pozo-Cruz B, Olivares PR, Parra Ja. Balance training reduces fear of falling and improves dynamic balance and isometric strength in institutionalised older people: a randomised trial. *J Physiother* 2012;58(2):97–104.
- [16] Dal Bello-Haas V, Klassen L, Sheppard MS, Metcalfe A. Psychometric properties of activity, self-efficacy, and quality-of-life measures in individuals with Parkinson disease. *Physiother Can* 2011;63(1):47–57.
- [17] Deathe AB, Miller WC. The L test of functional mobility: measurement properties of a modified version of the timed up & go test designed for people with lower-limb amputations. *Phys Ther* 2005;85(7):626–35.
- [18] Kim JS, Chu DY, Jeon HS. Reliability and validity of the L test in participants with chronic stroke. *Physiotherapy* 2015;101(2):161–5.
- [19] Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. *Neurology* 1967;17(5):427–42.
- [20] Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86(2):420–8.
- [21] Fleiss JL. The measurement of inter-rater agreement. In: *Statistical methods for rates and proportions*. New York: John Wiley; 1981.
- [22] Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. *Phys Ther* 2006;86(5):735–43.
- [23] Bland JM, Altman DG. Statistical method for assessing agreement between two methods of clinical measurement. *Lancet* 1986;307–10.
- [24] Prion S, Haerling KA. Making sense of methods and measurement: Pearson product-moment correlation coefficient. *Clin Simul Nurs* 2014;10(11):587–8.
- [25] The Australia and New Zealand Melanoma Trials Group. Statistical decision tree wizard [Internet]. The Australian and New Zealand Melanoma Trials Group; 2017.
- [26] Verheyden G. Psychometric properties of 3 functional mobility tests for people with Parkinson disease. *Am Phys Ther Assoc* 2014;94(8):1135–43.
- [27] Brusse KJ, Zimdars S, Zalewski KR, Steffen TM. Testing functional performance in people with Parkinson disease. *Phys Ther* 2005;85(2):134–41.

Available online at www.sciencedirect.com

ScienceDirect