

ASSESSING TIMED UP AND GO IN PARKINSON'S DISEASE: RELIABILITY AND VALIDITY OF TIMED UP AND GO ASSESSMENT OF BIOMECHANICAL STRATEGIES

Bruna Adriana DA SILVA, PT, MSc^{1,2}, Christina D. C. M. FARIA, PT, PhD³, Mariana Palla SANTOS, PT, MSc^{1,2} and Alessandra SWAROWSKY, PT, PhD^{1,2}

From ¹Brazilian Parkinson's disease Rehabilitation Initiative (BPaRkI), ²Physical Therapy Postgraduate Program, Physical Therapy Department, Santa Catarina State University (UDESC), Florianópolis and ³Physical Therapy Department, Federal University of Minas Gerais (UFMG), Belo Horizonte, Brazil

Objective: To investigate the measurement properties of the Timed Up and Go Assessment of Biomechanical Strategies (TUG-ABS) to determine its adequacy for use with individuals with Parkinson's disease.

Subjects: Fifty individuals with Parkinson's disease.

Design: Diagnostic accuracy.

Methods: The study investigated the following properties: reliability (inter-examiner, intra-examiner, test-retest, internal consistency and minimal detectable change), construct validity, and floor and ceiling effect.

Results: Considering the total score, the inter-examiner, test-retest and intra-examiner reliabilities were classified as excellent ($0.95 \leq \text{intra-class correlation coefficient (ICC)} \leq 0.99$). The TUG-ABS presented excellent internal consistency ($\alpha = 0.98$). The minimal detectable change was 3.82 points. The construct validity between the TUG-ABS and the Unified Parkinson's Disease Rating Scale (UPDRS) – part III was classified as moderate ($\rho = -0.62$). Significant, elevated and positive correlations were obtained between TUG-ABS and the Balance Evaluation System Test (BESTest)-VI ($\rho = 0.72$) and negative correlations between TUG-ABS and TUG ($\rho = -0.78$). The discriminant function obtained with the total score of TUG-ABS classified 60% of the individuals correctly with respect to the group (determined by the performance in TUG) to which they belonged. One-way analysis of variance (ANOVA) showed that TUG-ABS discriminated the individuals with Parkinson's disease in all stages according to Hoehn & Yahr. There was a ceiling effect of 22%.

Conclusion: TUG-ABS presented adequate measurement properties in individuals with Parkinson's disease.

Key words: validity; reliability; Parkinson's disease; physiotherapy; sensitivity; specificity.

Accepted May 19, 2017; Epub ahead of print Sep 26, 2017

J Rehabil Med 2017; 49: 00–00

Correspondence address: Alessandra Swarowsky, Brazilian Parkinson's disease Rehabilitation Initiative (BPaRkI), Physical Therapy Postgraduate Program, Physical Therapy Department, Santa Catarina State University (UDESC), Florianópolis, Brazil. Rua Pascoal Simone, 358. Zip code 88080350, Brazil. E-mail: alessandra.martin@udesc.br

Mobility changes are an inevitable consequence of Parkinson's disease (PD) and an important

cause of reduced quality of life in this population (1). Functional mobility is defined as the ability of a person to move in an independent and safe way in diverse environments and carry out everyday activities, such as locomotion and postural transfers (2). In PD the postural biomechanical changes related to progression of the disease contribute greatly to compromising functional mobility (3). Amongst these changes, the most marked are a narrow base and forward inclination of the trunk, internal rotation of the shoulders, and flexion of the hips and knees (3–6). These changes are reflected in an increase in flexor tonus and weakness of the extensor muscles, and can be associated with common negative effects on gait ability, such as reduced speed, decreased pace length and decreased arm swing (3–6).

Of the tests available to evaluate compromised functional mobility in PD, the Timed Up and Go test (TUG) is one of the most used, reported and highly recommended (7–9). TUG was defined by Isles et al. (10) as a measure of the time taken by an individual to carry out some functional manoeuvres, such as getting up, gait, turning and sitting down. Although commonly used and recommended, the outcome provided by TUG is the time of execution, which is of limited value for diagnoses or planning of treatment, since it fails to identify what is compromised in carrying out the task. Thus, Timed Up and Go Assessment of Biomechanical Strategies (TUG-ABS) was developed to be able to identify biomechanical strategies while carrying out TUG, and its measurement properties have been investigated for hemiparetic individuals (11). The outcome provided by TUG-ABS completes the time measure obtained using TUG. With TUG-ABS, in addition to obtaining a predictive measure of falls for the individuals, one can add standardized information about biomechanical strategies adopted by these individuals while carrying out the test (11–13). This information could improve diagnosis related to mobility limitation, enabling better clinical reasoning and treatment planning.

The aim of this study was therefore to investigate two principal domains of measurement properties of the TUG-ABS (reliability and validity, comparing accuracy between the tests) in order to determine whether the instrument is adequate for use in clinical practice in individuals with PD. We hypothesized that the TUG-ABS would be reliable and valid for

assessment of biomechanical strategies in individuals with PD, during getting up, turning and sitting down performed in sequence.

METHODS

This prospective study aimed to investigate the measurement properties of the TUG-ABS in 2 principal domains (reliability and validity) in individuals with PD. Both the planning of the investigation and the analysis of the measurement properties followed the recommendations of COSMIN (Consensus-based Standards Measurements Instruments – checklist) (14–16).

The research project was submitted to the Ethics Committee for Research using Human Beings of the Santa Catarina State University (UDESC, Florianópolis, Santa Catarina, Brazil) according to the terms of Resolution 466/2012, and was approved under report number 1.251.231. All participants were informed about the study objectives and signed a Free and Clarified Term of Consent (FCTC).

Participants

Fifty individuals with PD were included, according to the recommendations of COSMIN with respect to the number of participating individuals for a good study (14–16). They were selected by convenience, were of both sexes, and were all participants in the extension project “Neurofunctional Rehabilitation in Parkinson’s Disease” carried out at the Catarinense Rehabilitation Center of the State Secretary of Health of Santa Catarina, Brazil. All evaluations were carried out in the “on” medication phase, with intervals of between 2 and 4 h after administration and always in the afternoon period.

Inclusion criteria were: individuals diagnosed with PD and confirmed by a neurologist according to UK Brain Bank criteria (17); between stages 1 and 4 of the disease (Hoehn & Yahr) (18); cognitive level in the Mini Mental State Examination (MMSE) (19) relative to education levels given as cut-off points by Bertolucci et al. (20) (score ≥ 13 for illiterate subjects, ≥ 18 for low/medium education level, and ≥ 26 for a high education level); and on stable medication (for at least 4 weeks). Exclusion criteria were: individuals presenting other associated neurological or orthopaedic diseases and/or joint limitations that compromise the musculoskeletal system making mobility difficult, or important and/or severe dyskinesia (greater than 2 in item 33 of the UPDRS) (21).

Procedures

Participants’ sociodemographic data and clinical history were obtained with an evaluation form, developed previously for this purpose. All evaluation instruments were applied in the following order: Mini Mental State Examination (MMSE) (19), Hoehn & Yahr stage scale (18), and the motor section of the UPDRS–part III (21). The TUG test (22); section VI of the Balance Evaluation System Test (BESTest) (23) and the TUG-ABS test (11) were applied in sequence.

The TUG-ABS test is composed of 15 items that allow evaluation of the movement strategies adopted to carry out TUG, subdivided into the test activities: getting up from a seated position, gait, turning round 180°, and sitting down from a standing position. The total score varies from 15 to 45, with higher score indicating better performance by the individual (11–13).

Standardization of investigation of the measurement properties

Two evaluators were submitted to a 2-week training period to familiarize them with the TUG-ABS instrument, following orientation by the authors (11, 12). The inter-examiner, test-retest and intra-examiner reliabilities were evaluated. The first was established by comparison of the results of evaluations carried out by 2 independent evaluators (EV1 and EV2), carried out on the same day with a 1-h interval between them, the evaluators having no contact with each other during or in the intervals of the evaluation. The order of evaluations was always EV1 first followed by EV2. The second consisted of 2 evaluations carried out by the same evaluator (EV1) with a 1-week interval. Both evaluators had 5 years of experience in the evaluation of PD. Intra-examiner reliability was also investigated using a video, according to the procedures adopted by the author of the instrument (12): 3 Casio EX-FH20® video cameras were used to record the performance in TUG by the participants, while the evaluator (EV1) gave the scores for TUG-ABS from observations in real time. The TUG test was carried out according to the protocol recommended by Podsiadlo & Richardson (22). Four weeks later the same evaluator (EV1) gave his scores for the TUG-ABS again from the previously recorded videos of the same evaluations in random order. The videos were processed and edited using Adobe Premiere Pro 2.0® software, which allows for the 3 vision points to be grouped in the same file and observed simultaneously on screen. The cameras were positioned as shown in Fig. 1. The camera focus was positioned at a height of 105 cm. All evaluations were carried out according to the procedures adopted for the development of TUG-ABS (24).

Statistical analysis

Sample characterization. The sociodemographic characteristics of the participants were described using descriptive statistics (mean and standard deviation (SD)).

Floor and ceiling effects. The floor and ceiling effects were verified by calculating the percentage of individuals who obtained the minimum and maximum scores in the TUG-ABS. When the incidence of either of these scores was greater than 15% of the sample, the floor (minimum) and/or ceiling (maximum) effects were present, and this criterion was adopted in the present study (25).

Psychometric domains: reliability and validity. The intra-class correlation coefficient (ICC) with 95% confidence interval (95% CI) was used to evaluate the inter-examiner, intra-examiner and test-retest reliabilities. In the presence of significant correlation ($\alpha=5\%$), the following classification was adopted: weak

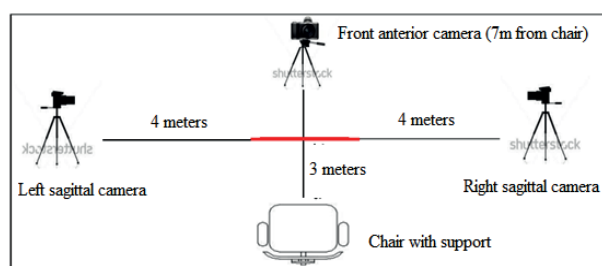


Fig. 1. Camera positioning scheme to register performance in Timed Up and Go test.

agreement, $ICC < 0.40$; moderate agreement, $ICC \leq 0.75$; and excellent agreement, $ICC > 0.75$ (24).

The inter-examiner reliability was also evaluated for the scoring of each item of the TUG-ABS using the weighted kappa (k_p) for the individual items. In the presence of significant agreement ($\alpha = 5\%$), the following classification was adopted: excellent for k_p values from 0.81 to 1.0; substantial, 0.61–0.80; moderate, 0.41–0.60; considerable, 0.21–0.40; and slight, 0–0.20 (27). A Bland & Altman plot analysis was used to make a more detailed analysis of eventual differences between the scores of the evaluators in the TUG-ABS test (25).

The internal consistency, another type of reliability, was established using Cronbach's alpha coefficient, adopting a value between 0.70 and 0.90 to classify good agreement between the items of TUG-ABS (26).

The minimal detectable change (MDC) was calculated from the confidence interval, standard deviation of the base evaluation (EV1) and the correlation value of the test-retest, referring to the minimal amount of change that is not due to variation on the TUG-ABS scale. The MDC was measured using the formula $MDC = Z_{score} \times \sqrt{2[1 - r_{test-retest}]}$, considering a 95% CI (25).

We hypothesized that the internal consistency of TUG-ABS would vary between 0.7 and 0.9, and the reliabilities should exceed 0.75 for ICC and 0.6 for k_p value (26, 27).

Construct validity was investigated from the convergence analysis between the scores in TUG-ABS with the results of TUG; between the scores in TUG-ABS and the total score of UPDRS-III, and between the scores in TUG-ABS and the total score of section VI of the BESTest, which refers to gait stability. Non-parametric Spearman's correlation was used for these analyses. In the presence of significant correlation ($\alpha = 5\%$), the following classification was adopted: low $0.20 \leq \rho \leq 0.49$, moderate $0.50 \leq \rho \leq 0.69$, high $0.7 \leq \rho \leq 0.89$ and very high $0.90 \leq \rho \leq 1$ (28). We hypothesized that, for construct validity for convergence analysis, we expected that the correlation between the UPDRS-III and TUG with TUG-ABS was negative and higher than 0.5, and with BESTest section IV and TUG-ABS, positive and higher than 0.5.

The construct validity was also investigated by discriminant analysis, a multivariate analysis that enables evaluation of how the variables differ between the groups and how the individuals are classified within the group they belong to from these variables. For this analysis, the individuals with PD were divided into 3 subgroups based on their performance in the TUG test (quick, moderate or slow). For this purpose, the TUGs of each individual were organized in increasing order, as already carried out in an earlier study with another population group (24). The inter-tertile interval, with decomposition points of 33% and 66%, was then used to divide the individuals within the 3 above-mentioned subgroups. One-way analysis of variance (ANOVA) was used to compare the TUG times of the 3 subgroups, followed by the *post-hoc* test ($\alpha = 5\%$), to determine if the subgroups differed with respect to the TUG time (24, 27). This first analysis with ANOVA was important to add information about the construct validity of the TUG-ABS, since the development of this instrument was based on the selection of variables that could differentiate the individuals who carried out the TUG test in different times.

Discriminant analysis can be used for descriptive and predictive purposes. In the present study, discriminant analysis was carried out to investigate whether the TUG-ABS scores could predict the association between the groups determined by their performance in TUG (quick, moderate and slow) (24, 29). We

hypothesized that TUG-ABS could classify correctly the groups determined by TUG performance (quick, moderate and slow).

Finally, comparison of the known groups was also used to investigate the construct validity of the TUG-ABS (to verify whether the groups determined by Hoehn & Yahr are the same as those determined by the TUG-ABS). For this purpose the known groups were determined according to the Hoehn & Yahr scale using 3 degrees of impairment: mild (stages 1 and 2), moderate (stage 3) and severe (stage 4). One-way ANOVA with Tukey's *post-hoc* ($\alpha = 5\%$) was used to compare the TUG-ABS scores obtained by each of these 3 groups (to verify differences between the 3 groups). We hypothesized that TUG-ABS would also be able to differentiate the groups determined by Hoehn & Yahr.

All the data were analysed using the MedCalc 12.5.0 and SPSS 20.0 programs, both from Windows, accepting a significance level of 5% for all these procedures.

RESULTS

The present study was carried out from August 2015 to July 2016.

Sample characterization

The mean age of the individuals with PD who took part in this study was 67.38 years (SD 8.98), with no sex predominance. According to the Hoehn & Yahr scale there was a predominance of participants in the mild (1 & 2) and moderate (3) stages, of 82%. Table I shows the descriptive statistics of the clinical-demographic characteristics of the sample. No adverse events occurred during the tests.

Table I. Clinical and sociodemographic characteristics of the study participants

Characteristics		Variation (range)
Sex, n (%)		
Male	25 (50)	
Female	25 (50)	
Age, (years), mean (SD)	67.4 (9.0)	46–87
Diagnosis time, (years), mean (SD)	8.6 (7)	1–30
Stage-H&Y, n (%)		
1 (mild)	5 (10)	
2 (mild)	15 (30)	
3 (moderate)	21 (42)	
4 (severe)	9 (18)	
MMSE (0–30), mean (SD)	24.6 (3.9)	16–30
UPDRS – part III (0–56), mean (SD)	14.0 (6.9)	3–30
BESTest – section VI (0–21), mean (SD)	13.7 (4.6)	5–21
TUG (s), mean (SD)	16.7 (9.2)	7–53
TUG-ABS (15–45), mean (SD)	37.3 (6.8)	20–45
Medication in use, n (%)		
Amantadine	15 (30)	
Biperiden	4 (8)	
Entacapone	5 (10)	
Levodopa	46 (92)	
Pramipexole	13 (26)	

H&Y: Hoehn & Yahr; MMSE: Mini Mental State Examination; SD: standard deviation; UPDRS: Unified Parkinson's Disease Rating Scale; BESTest: Balance Evaluation System Test; TUG: Timed Up and Go test; TUG-ABS: Timed Up and Go Assessment of Biomechanical Strategies.

Floor and ceiling effects

The TUG-ABS test presented a ceiling effect of 22% (11/50) and 10 of these 11 individuals were classified as being in the mild impairment stage (25).

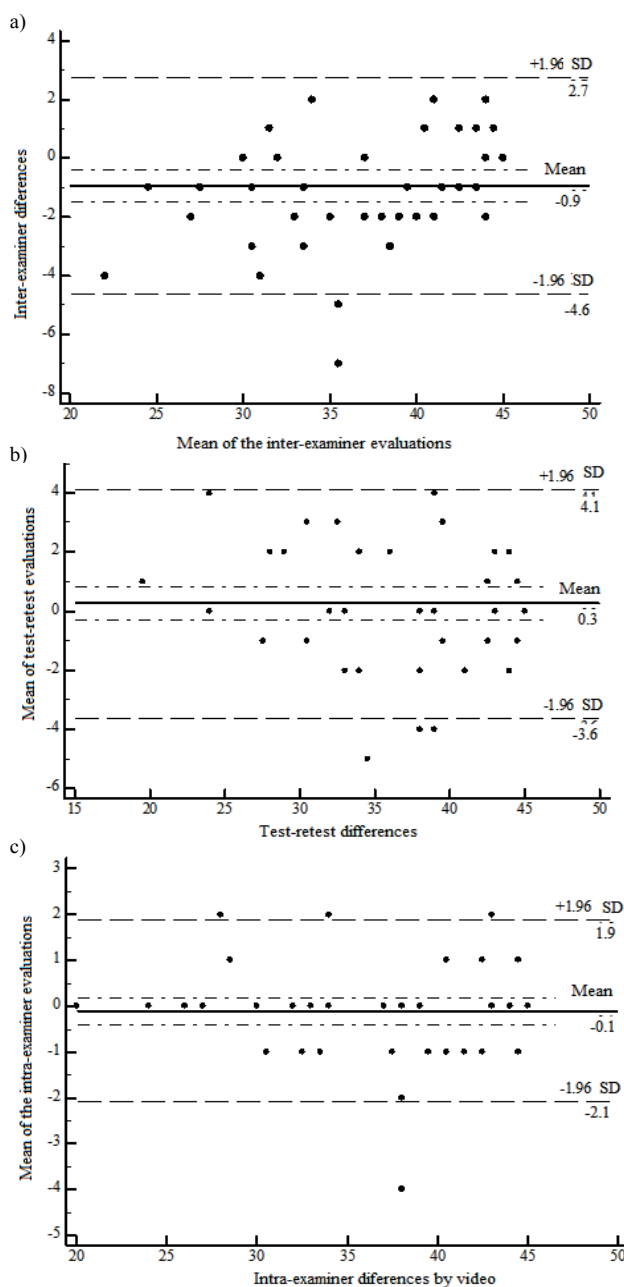


Fig. 2. Bland & Altman plot analysis showing inter-evaluator and test-retest agreement considering the total Timed Up and Go Assessment of Biomechanical Strategies (TUG-ABS) scores. (a) Mean of the evaluations and inter-examiner differences that did not differ significantly from zero; (b) Mean of the evaluations and test-retest differences that also did not differ significantly from zero; (c) Mean of the evaluations and intra-examiner differences that also did not differ significantly from zero.

Reliability

In the evaluation of the inter-examiner reliability for the TUG-ABS test, the ICC (95% CI) for the total score was 0.95 (0.93–0.98). The value for k_p for the total score was 0.8 (0.74–0.86) and for the individual items (95% CI) varied between 0.27 and 0.73 (Table II). The minimum and maximum values for k_p refer to questions 7 and 12, respectively. The Bland & Altman plot (Fig. 2a) illustrates the inter-examiner agreement, considering the total scores. For the TUG-ABS test, the mean difference between the 2 evaluations did not differ significantly from zero, and the limits of agreement represented 6% and 10.2% of the variation of the instrument.

With respect to the test-retest reliability of TUG-ABS, the ICC (95% CI) for the total score was equal to 0.96 (0.93–0.98) and for the individual items varied between 0.39 (question 2) and 0.95 (question 4) (Table III). Analysis of the Bland & Altman plot (Fig. 2b) showed that the limits of agreement represented 8% and 9.1% of the scale variation.

In the intra-examiner reliability according to the TUG-ABS video, the ICC (95% CI) for the individual items varied between 0.56 (question 2) and 0.97 (question 12), and was equal to 0.99 (0.98–0.99) for the total score (Table IV). Analysis of the Bland & Altman plot (Fig. 2c) showed that the limits of agreement represented 4.2% and 4.6% of the variation of the instrument.

The internal consistency of the TUG-ABS test was considered excellent, with a Cronbach's alpha coefficient of 0.98 ($\alpha=0.98$).

According to the MDC the total score of the TUG-ABS scale was 3.82 points.

Validity

With respect to the construct validity according to the convergence analysis, the total score of the TUG-ABS test presented a highly negative correlation ($p=-0.78$; $p<0.001$) with the TUG test, a moderately negative correlation with the UPDRS – part III ($p=-0.62$; $p<0.001$) and a highly positive correlation with section VI of the BESTest ($p=0.72$; $p<0.001$).

With respect to the construct validity investigated by discriminant analysis, the individuals with PD divided into 3 groups (quick, moderate and slow performance) presented significant differences in the TUG test (Table V): the 1-way ANOVA showed significant differences between the groups ($F(2)=31.01$; $p<0.001$) and the *post-hoc* showed significant differences between all 3 groups ($p<0.05$). As can be seen in Table VI, 76.5% of the individuals of the fast group, 18.8% of the moderate group and 82.4% of the slow group were correctly classified by the TUG-ABS score for the initially grouped

Table II. Inter-examiner reliability analysis for Timed Up and Go Assessment of Biomechanical Strategies (TUG-ABS)

Inter-examiners	Evaluator 1		Evaluator 2		kw value (95% CI)
	Median (min-max)	Mean (SD)	Median (min-max)	Mean (SD)	
<i>From seated to standing</i>					
Support of upper limb(s) associated with lateral flexion and/or trunk rotation	3 (2-3)	2.8 (0.4)	3 (1-3)	2.7 (0.5)	0.64 (0.41-0.86)
Attempts to stand up associated with strategy of sitting on edge of chair	3 (2-3)	2.86 (0.35)	3 (1-3)	2.76 (0.51)	0.44 (0.2-0.7)
Momentum generated by the first front trunk flexion and by extension of the trunk and lower limbs	2 (2-3)	2.48 (0.5)	3 (2-3)	2.66 (0.48)	0.56 (0.35-0.77)
<i>Gait</i>					
Step symmetry and length (majority of steps)	1 (1-3)	1.86 (0.99)	2.5 (1-3)	2.1 (0.95)	0.72 (0.55-0.89)
Initial feet contact with heels (majority of steps)	3 (1-3)	2.3 (0.93)	3 (1-3)	2.58 (0.81)	0.56 (0.4-0.79)
Hip extension in support phase: Backward positioning of thigh in relation to pelvis (majority of steps)	3 (1-3)	2.48 (0.88)	3 (1-3)	2.42 (0.90)	0.57 (0.32-0.82)
Balance phase-absence of foot contact with ground (majority of steps)	3 (1-3)	2.68 (0.74)	3 (1-3)	2.38 (0.92)	0.27 (0.01-0.55)
Forward progression of lower limbs without atypical trunk movement (majority of steps)	3 (1-3)	2.58 (0.78)	3 (1-3)	2.8 (0.6)	0.44 (0.13-0.75)
<i>Turning round</i>					
Ratio between the external and internal foot parts and the circumference in turning round	2 (1-3)	2.14 (0.75)	2 (1-3)	2.3 (0.76)	0.46 (0.27-0.65)
Number of steps taken to turn round (does not include steps taken pre- and post- turning round during gait)	3 (1-3)	2.54 (0.76)	3 (1-3)	2.74 (0.6)	0.59 (0.33-0.85)
Body rotation to complete the change in direction on turning round	3 (1-3)	2.64 (0.56)	3 (1-3)	2.68 (0.62)	0.46 (0.2-0.71)
Gait-turn round-gait sequence	3 (1-3)	2.58 (0.57)	3 (2-3)	2.56 (0.5)	0.73 (0.57-0.9)
<i>From standing to sitting</i>					
Sequence between gait, turning round to sit down and sitting down	3 (1-3)	2.52 (0.70)	2.5 (1-3)	2.46 (0.58)	0.6 (0.44-0.76)
Sequence and control as the pelvis and trunk approach the chair	2 (1-3)	2.22 (0.81)	2 (1-3)	2.4 (0.64)	0.46 (0.25-0.6)
Positioning of lower limbs and active knee flexion on sitting down on a chair	3 (1-3)	2.66 (0.52)	3 (1-3)	2.74 (0.52)	0.47 (0.27-0.68)
Total	38.5 (20-45)	37.34 (6.89)	40 (24-45)	38.28 (6.08)	0.8 (0.74-0.86)**

** $p < 0.01$; 95% CI: 95% confidence interval; (min-max), minimum and maximum values; SD: standard deviation; kw: weighted kappa.

Table III. Test-retest reliability analysis for Timed Up and Go Assessment of Biomechanical Strategies (TUG-ABS)

Test-retest	Evaluator 1 (EV1)		Evaluator 1 (EV2)		ICC (95% CI)
	Median (min-max)	Mean (SD)	Median (min-max)	Mean (SD)	
<i>From seated to standing</i>					
Support of upper limb(s) associated with lateral flexion and/or trunk rotation	3 (2-3)	2.8 (0.4)	3 (1-3)	2.64 (0.66)	0.50 (0.26-0.68)
Attempts to stand up associated with strategy of sitting on edge of chair	3 (2-3)	2.86 (0.35)	3 (1-3)	2.84 (0.42)	0.39 (0.13-0.6)
Momentum generated by the first front trunk flexion and by extension of the trunk and lower limbs	2 (2-3)	2.48 (0.5)	3 (1-3)	2.52 (0.54)	0.70 (0.53-0.82)**
<i>Gait</i>					
Step symmetry and length (majority of steps)	1 (1-3)	1.86 (0.99)	1 (1-3)	1.88 (1)	0.95 (0.91-0.97) **
Initial feet contact with heels (majority of steps)	3 (1-3)	2.3 (0.93)	3 (1-3)	2.28 (0.97)	0.65 (0.45-0.78)
Hip extension in support phase:			3 (1-3)	2.28 (0.97)	
Backward positioning of thigh in relation to pelvis (majority of steps)	3 (1-3)	2.48 (0.88)			0.78 (0.65-0.87)**
Balance phase-absence of foot contact with ground (majority of steps)	3 (1-3)	2.68 (0.74)	3 (1-3)	2.74 (0.66)	0.65 (0.46-0.78)
Forward progression of lower limbs without atypical trunk movement (majority of steps)	3 (1-3)	2.58 (0.78)	3 (1-3)	2.46 (0.84)	0.76 (0.62-0.86)**
<i>Turning round</i>					
Ratio between the external and internal foot parts and the circumference in turning round	2 (1-3)	2.14 (0.75)	2 (1-3)	2.16 (0.73)	0.54 (0.1-0.71)
Number of steps taken to turn round (does not include steps taken pre and post turning round during gait)	3 (1-3)	2.54 (0.76)	3 (1-3)	2.56 (0.61)	0.72 (0.56-0.83)**
Body rotation to complete the change in direction on turning round	3 (1-3)	2.64 (0.56)	3 (1-3)	2.68 (0.55)	0.74 (0.58-0.84)**
Gait-turn round-gait sequence	3 (1-3)	2.58 (0.57)	3 (1-3)	2.52 (0.61)	0.86 (0.77-0.92)**
<i>Standing to sitting</i>					
Sequence between gait, turning round to sit down and sitting down	3 (1-3)	2.52 (0.70)	3 (1-3)	2.54 (0.61)	0.65 (0.46-0.78)
Sequence and control as the pelvis and trunk approach the chair	2 (1-3)	2.22 (0.81)	3 (1-3)	2.48 (0.67)	0.68 (0.5-0.8)
Positioning of lower limbs and active knee flexion on sitting down on a chair	3 (1-3)	2.66 (0.52)	3 (1-3)	2.5 (0.64)	0.5 (0.26-0.68)
Total	38.5 (20-45)	37.34 (6.89)	39 (19-45)	37.11 (7.28)	0.96 (0.93-0.98)**

** $p < 0.01$; 95% CI: 95% confidence interval; ICC: intraclass correlation coefficient; (min-max), minimum & maximum values; EV1: first evaluation; EV2: second evaluation; SD: standard deviation.

Table IV. Intra-examiner reliability analysis by video for Timed Up and Go Assessment of Biomechanical Strategies (TUG-ABS)

	Evaluator 1 (EVV)		Evaluator 1 (EV1)		ICC (CI 95%)
	Median (min-max)	Mean (SD)	Median (min-max)	Mean (SD)	
Intra-examiner					
<i>From seated to standing</i>					
Support of upper limb(s) associated with lateral flexion and/or trunk rotation	3 (1-3)	2.7 (0.54)	3 (2-3)	2.8 (0.4)	0.62 (0.42-0.76)
Attempts to stand up associated with strategy of sitting on edge of chair	3 (2-3)	2.88 (0.33)	3 (2-3)	2.86 (0.35)	0.56 (0.33-0.72)
Momentum generated by the first front trunk flexion and by extension of the trunk and lower limbs	3 (1-3)	2.56 (0.54)	2 (2-3)	2.48 (0.5)	0.79 (0.65-0.87)**
<i>Gait</i>					
Step symmetry and length (majority of steps)	1.5 (1-3)	1.96 (0.99)	1 (1-3)	1.86 (0.99)	0.91 (0.85-0.95)**
Initial feet contact with heels (majority of steps)	3 (1-3)	2.26 (0.96)	3 (1-3)	2.3 (0.93)	0.79 (0.66-0.88)**
Hip extension in support phase:					
Backward positioning of thigh in relation to pelvis (majority of steps)	3 (1-3)	2.3 (0.95)	3 (1-3)	2.48 (0.88)	0.76 (0.62-0.86)**
Balance phase-absence of foot contact with ground (majority of steps)	3 (1-3)	2.58 (0.81)	3 (1-3)	2.68 (0.74)	0.85 (0.76-0.91)**
Forward progression of lower limbs without atypical trunk movement (majority of steps)	3 (1-3)	2.58 (0.78)	3 (1-3)	2.58 (0.78)	0.83 (0.72-0.9)**
<i>Turning round</i>					
Ratio between the external and internal foot parts and the circumference in turning round	2 (1-3)	2.24 (0.77)	2 (1-3)	2.14 (0.75)	0.74 (0.59-0.85)**
Number of steps taken to turn round (does not include steps taken pre and post turning round during gait)	3 (1-3)	2.64 (0.6)	3 (1-3)	2.54 (0.76)	0.81 (0.69-0.89)**
Body rotation to complete the change in direction on turning round	3 (1-3)	2.66 (0.55)	3 (1-3)	2.64 (0.56)	0.84 (0.73-0.9)**
Gait-turn round-gait sequence	3 (1-3)	2.56 (0.57)	3 (1-3)	2.58 (0.57)	0.97 (0.94-0.98)**
<i>Standing to sitting</i>					
Sequence between gait, turning round to sit down and sitting down	3 (1-3)	2.52 (0.61)	3 (1-3)	2.52 (0.70)	0.86 (0.76-0.92)**
Sequence and control as the pelvis and trunk approach the chair	3 (1-3)	2.44 (0.73)	2 (1-3)	2.22 (0.81)	0.75 (0.6-0.85)**
Positioning of lower limbs and active knee flexion on sitting down on a chair	3 (1-3)	2.56 (0.61)	3 (1-3)	2.66 (0.52)	0.66 (0.48-0.79)
Total	38.5 (20-45)	37.44 (6.99)	39.5 (20-45)	37.34 (6.89)	0.99 (0.98-0.99)**

** $p < 0.01$; 95% CI: 95% confidence interval; ICC: intraclass correlation coefficient; (min-max), minimum & maximum values; EVV: evaluation by video; EV1: first evaluation; SD: standard deviation.

cases. The discriminant function obtained with the total TUG-ABS score showed a canonical correlation of 0.740 (Wilks $\lambda = 0.452$, $\chi^2 = 37.308$, $p < 0.001$), with 60% of individuals correctly classified into the group to which they belonged.

With respect to the construct validity determined by comparison of known groups using the Hoehn & Yahr scale (Table V), the 1-way ANOVA showed significant differences between the known groups ($F(2) = 30.16$; $p < 0.05$), and using Tukey's *post-hoc* test it was observed that the TUG-ABS score differed significantly between all the groups: $p = 0.0001$ between the mild and moderate impairment stages; $p = 0.003$ between the moderate and severe impairment stages; and $p = 0.0001$ between the mild and severe impairment stages, where the higher the total score in the TUG-ABS, the less severe the impairment.

DISCUSSION

The aim of the study was to evaluate 2 measurement property domains of the TUG-ABS test in individuals with PD, which showed adequate results for reliability (intra- and inter-examiner, test-retest and internal con-

Table VI. Results of the discriminant analysis by know groups ($n = 50$)

True group	Fast n (%)	Moderate n (%)	Slow n (%)
Fast performance ($n = 17$)	13 (76.5)	3 (17.6)	1 (5.9)
Moderate performance ($n = 16$)	9 (56.3)	3 (18.8)	4 (25)
Slow performance ($n = 17$)	0	3 (17.6)	14 (82.4)
Total of the true group correctly classified: 30 (60)			

TUG: Timed Up and Go test; TUG-ABS: Timed Up and Go Assessment of Biomechanical Strategies; The bold numbers represent the individuals originally grouped together, who were classified correctly by the predictive variable.

Table V. Discriminant validity of known groups between the total score for the Timed Up and Go Assessment of Biomechanical Strategies (TUG-ABS) and the Hoehn & Yahr stage ($n = 50$)

Groups	n	TUG (SD) [variation]	CI (95%)	Mean of total score in TUG-ABS (SD) [variation]	95% CI	p -value
Fast performance	17	10.00 (1.33) [7.0-12.0]	9.32-10.69	-	-	<0.01
Moderate performance	16	13.88 (1.54) [12.25-17.0]	13.05-14.7	-	-	<0.01
Slow performance	17	26.13 (10.47) [17.10-53.0]	20.75-31.52	-	-	<0.01
H&Y 1 & 2 (mild)	20	11.04 (2.22) [7-15.8]	10.07-12.01	42.95 (3.12) [32-45]	41.49-44.41	<0.01
H&Y (moderate)	21	17.29 (9.19) [9.4-53]	13.2-21.2	35.52 (6.21) [20-45]	32.69-38.35	<0.01
H&Y 4 (severe)	9	28.04 (8.89) [20-49.6]	22.23-33.85	29.11 (2.8) [24-34]	26.95-31.26	<0.01

* $p < 0.05$; TUG: Timed Up and Go test; SD: standard deviation; H&Y: Hoehn & Yahr; 95% CI: 95% confidence interval.

sistency, the MDC also being determined and classified as excellent), and for validity (construct), when used in these individuals.

There was a ceiling effect of 22% for the TUG-ABS test; that is, more than 15% of the participants obtained the maximum total score of 45 points, showing the best performance, and, of these 22%, 91% were classified in the mild impairment stage. Although the instrument was sufficiently sensitive to evaluate individuals with PD in all stages of the disease, there was a ceiling effect for individuals showing the best performance (normally corresponding to the mild impairment stage) apparently signifying that the TUG-ABS test was not sensitive enough to capture alterations presented at the start of the disease. Another possible explanation for these results could be that in the mild stage, the common alterations in functional mobility were not sufficient to alter the biomechanical strategies adopted during the performance of TUG.

The inter-examiner reliability for the total TUG-ABS score presented excellent agreement ($ICC=0.95$), a similar result to that found for individuals with stroke (13). When the TUG-ABS items were evaluated for their kp values (0.27–0.73), it was noted that the lowest value referred to the question “Balance phase—absence of foot contact with the ground (majority of the steps)” with respect to the “Gait” phase of TUG. The present results corroborated with those found in the original version, with a $kp=0.24$ for the same task in post-stroke individuals (12).

The test-retest reliability was considered excellent for the total TUG-ABS score ($ICC=0.96$). When the TUG-ABS items were analysed individually (0.39–0.95), the item showing the least reliability was the question referring to “Attempts to get up from a seated position associated with the use of the strategy of sitting as close as possible to the edge of the chair” in the “Seated to on foot” phase of TUG.

In the intra-examiner reliability, the total TUG-ABS score showed excellent agreement ($ICC=0.99$) and, once again, the lowest reliability referred to the question “Attempts to get up from a seated position associated with the use of the strategy of sitting as close as possible to the edge of the chair”. The item showing the greatest agreement was the same for the 3 reliabilities tested: “gait-turn round-gait sequence” in the “Turning round” phase of TUG, not agreeing with the results of the original version which presented the item “Attempts to get up from a seated position associated with the use of the strategy of sitting as close as possible to the edge of the chair” as the question with greater reliability (kp of 0.73 to 1.00). Reliability was shown to be excellent for both video analysis and application of the instrument in real time, the form of

evaluation remaining at the discretion of the physiotherapist according to his or her objectives. In relation to internal consistency, the TUG-ABS presented $\alpha=0.98$, this result being similar to that of the original version ($\alpha=0.87$), both being considered excellent.

The MDC of 3.82 points found for the total TUG-ABS score was considered excellent, representing the minimum amount of change necessary for the total score of the instrument to be considered a true change in the period of time between evaluations by the same evaluator (24, 25). Thus, the MDC obtained indicated that the change in the total TUG-ABS scores presented a chance of less than 8.48% to be due to random variation or measurement error (30), helping the physiotherapist and permitting a better interpretation of the results obtained with the use of the instrument at different times or with different evaluators.

When analysing the construct validity by convergence analysis, the result of the set of 3 analyses showed the adequacy of this property of the TUG-ABS in individuals with PD (27). The TUG-ABS showed a moderately negative correlation with UPDRS – part III. The items 13, 14, 15, 29 and 30 of the UPDRS, which evaluate postural and gait instability in PD, were recommended for their excellent internal consistency, although they showed a floor effect in the mild impairment stage and did not include items of gait performance. This could explain their moderate correlation with the TUG-ABS, which has items evaluating gait performance in its construct (9).

When the TUG-ABS was correlated with section VI of the BESTest, a high positive correlation was found. Bloem et al. (9) evaluated all the instruments used to evaluate posture, gait and equilibrium in PD, according to the orientations of COSMIN, and concluded that no instrument covered all the gait characteristics specific and relevant to PD. Thus they recommended the best evaluation instruments according to the various constructs of interest. These same authors recommended the use of the MiniBESTest and of the BESTest for the functional evaluation of gait and equilibrium in individuals with PD, since they cover an ample variety of constructs, although there is a need for extra materials to carry them out and a time of 10–15 min for their application (9). In the present study, section VI, related to gait stability was correlated with TUG-ABS, and high correlation was found, possibly because they refer to similar constructs.

TUG-ABS was developed to complement TUG, a measurement with adequate validity and reliability to evaluate functional mobility in individuals with PD (10, 22, 31, 33). When compared with the TUG time, the TUG-ABS score presented a highly negative correlation ($p=-0.78$), indicating that the shorter the TUG

time, the higher the TUG-ABS score, characterizing better performance. A time longer than 11.5 s in TUG is considered predictive of falls in PD (32), and the test is capable of discriminating PD fallers from non-fallers (33), as well as presenting a value for MDC of 3.5 s in these individuals (34). Since the only standardized measurement outcome for TUG is the time, which, although presenting a series of advantages, is limited by providing little information about the quality of the performance, the use of TUG-ABS complements the information obtained with TUG. The TUG-ABS provides information about the biomechanical strategies used by these individuals during the carrying out of important functional activities, such as those that constitute the TUG test, and this information is important for making clinical decisions with respect to these individuals. Thus TUG-ABS can be used to complement the results of TUG in individuals with PD, as already indicated for post-stroke individuals (12, 13).

The construct validity, analysed by discriminant analysis, showed a global classification precision of 60%, indicating that the predictions of the group members in relation to the TUG-ABS score were correctly classified in the majority of cases initially grouped considering the TUG results. Only the moderately impaired group was poorly classified (18.8%). Analysis of the confidence intervals of the individuals with moderate and slow performance (Table V) indicated a possible explanation for this result: an extensive time interval can be observed, and hence the individuals in the sample in the present study may not have contemplated the possible diversity in performance. Although the 3 groups presented statistically different TUG results, the groups were formed considering a sample selected by convenience. It is possible that, if more x with different performances in TUG had been evaluated, the formation of groups with different performances in TUG might have contemplated a greater variability in performance that could have been better identified by TUG-ABS.

In relation to the construct validity by comparison of known groups, it was verified that the groups in more advanced stages presented lower total scores in the TUG-ABS test, indicating greater impairment of the biomechanical strategies. The statistical analysis revealed that the TUG-ABS test (Table V) was capable of discriminating between the individuals with PD in all stages of the disease. However, these results should be considered with caution, since the severely impaired group consisted of a small sample (9 individuals). This discriminatory capacity of the degree of impairment by instruments is important, since the physiotherapist has difficulty in differentiating mild from moderate stages for questions that the Hoehn & Yahr scale (18) does not consider.

One of the limitations of the present study was the small number of participants in the severe stage of PD, and hence one cannot generalize the result for application in the population in more advanced stages of the disease. Further studies are required to ratify the results already found, including a greater variability in the individuals with respect to their performance in TUG, which could provide answers to other important questions raised by its use.

In conclusion, the TUG-ABS was shown to be an instrument with reliability and construct validity, with the accuracy to identify the biomechanical characteristics and strategies used by individuals with PD while carrying out the TUG test. The TUG-ABS is therefore of use in clinical practice.

REFERENCES

- Horak FB, Macini M. Objective biomarkers of balance and gait for Parkinson's disease using body-worn sensors. *Mov Disord* 2013; 28: 1544–1551.
- King L, Horak FB. Delaying mobility disability in people with Parkinson disease using a sensorimotor agility exercise program. *Phys Ther* 2009; 89: 384–393.
- Dionisio VC, Curtarelli MB, Souza LAP. Sitting movement in elderly subjects with and without Parkinson's disease: a biomechanical study. *J Electromyogr Kinesiol* 2013; 23: 948–957.
- Plotnik M, Giladi N, Dagan Y, Hausdorff FJM. Postural instability and fall risk in Parkinson's disease: impaired dual tasking, pacing, and bilateral coordination of gait during the 'ON' medication state. *Exp Brain Res* 2011; 210: 529–538.
- Carpenter MG, Bloem BR. Postural control in Parkinson patients: a proprioceptive problem? *Exp Neurol* 2011; 227: 26–30.
- Ziegler K, Schroeteler F, Baumann A, Fietzek U. A new rating instrument to assess festination and freezing. *Mov Disord* 2010; 25: 1012–1018.
- Keus SHJ, Munneke M, Graziano M, et al. European physiotherapy guideline for Parkinson's disease. 2014; KNGF/ParkinsonNet, the Netherlands. Available from: <http://www.parkinsonnet.info/guidelines/european-guidelines-overview>.
- Kegelmeyer D, Ellis T, Esposito A, Gallagher R, Harro CC, Hoder J, et al. The Parkinson evidence database to guide effectiveness (PDEDGE). *Neurology Section* 2014. Available from: www.neuropt.org/professional-resources/neurology-section-outcome-measures-recommendations/parkinson-disease.
- Bloem BR, Marinus J, Almeida Q, Dibble L, Nieuwboer A, Post B, et al. Measurement instruments to assess posture, gait, and balance in Parkinson's disease: critique and recommendations. *Mov Disord* 2016; 31: 1342–1355.
- Isles R, Choy NL, Steer M, Nitz JC, et al. Normal values of balance tests in women aged 20–80. *J Am Geriatr Soc* 2004; 52: 1367–1372.
- Faria CD, Teixeira-Salmela LF, Araújo PA, Polese JC, Nascimento LR, Nadeau S. TUG-ABS Portuguese-Brazil: a clinical instrument to assess mobility of hemiparetic subjects due to stroke. *Rev Neurocienc* 2015; 23: 357–367.
- Faria CD, Teixeira-Salmela LF, Nadeau S. Development and validation of an innovative tool for the assessment of biomechanical strategies: the timed "up and go"–assessment of biomechanical strategies (TUG-ABS) for individuals with stroke. *J Rehabil Med* 2013; 45: 232–240.
- Faria CD, Teixeira-Salmela LF, Nadeau S. Clinical testing

- of an innovative tool for the assessment of biomechanical strategies: the timed "up and go" - assessment of biomechanical strategies (TUG-ABS) for individuals with stroke. *J Rehabil Med* 2013; 45: 241–247.
14. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60: 34–42.
 15. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010; 19: 539–549.
 16. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; 63: 737–745.
 17. Hughes AJ, Daniel SE, Kilfor L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* 1992; 55: 4.
 18. Hoehn M, Yahr M. Parkinsonism: onset, progression and mortality. *Neurology* 1967; 17: 427–442.
 19. Folstein MF, Robins L, Helzer NJE. The mini-mental state examination. *Arch Gen Psychiatry* 1983; 40: 812.
 20. Bertolucci PH, Brucki SM, Campacci SR, Juliano Y. [The Mini-Mental State Examination in an outpatient population: influence of literacy.] *Arq Neuropsiquiatr* 1994; 52: 1–7 (in spanish).
 21. Movement Disorder Society. Task force on rating scales for Parkinson's disease the unified Parkinson's disease rating scale (UPDRS): status and recommendations. *Mov Disord* 2003; 18: 738–750.
 22. 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.
 23. Maia AC, Rodrigues-de-Paula F, Magalhães LC, Teixeira RL. Cross-cultural adaptation and analysis of the psychometric properties of the balance evaluation systems test and MiniBESTest in the elderly and individuals with Parkinson's disease: application of the Rasch model. *Braz J Phys Ther* 2013; 17: 195–217.
 24. Faria CD, Teixeira-Salmela LF, Nadeau S. Predicting levels of basic functional mobility, as assessed by the Timed "Up and Go" test, for individuals with stroke: discriminant analyses. *Disabil Rehabil* 2013; 35: 146–152.
 25. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. *Phys Ther* 2006; 86: 735–743.
 26. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159–174.
 27. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 2nd edition. New Jersey: Prentice-Hall; 2000.
 28. Munro B. Statistical methods for health care research. 5th edn. Philadelphia: Lippincott Williams and Wilkins; 2005.
 29. Tabachnick BG, Fidell LS. Using multivariate statistics. 3rd edition. New York: Harper Collins College Publishers; 1996.
 30. Stratford PW, Binkley J, Solomon P, Finch E, Gill C, Moreland J. Defining the minimum level of detectable change for the Roland-Morris questionnaire. *Phys Ther* 1996; 76: 359–365.
 31. Morris S, Morris E, Iansek R. Reliability of measurements obtained with the Timed "Up and Go" test in people with Parkinson's disease. *Phys Ther* 2001; 81: 810–817.
 32. Nocera JR, Stegemöller EL, Malaty IA, Okun MS, Marsiske M, Hass CJ. Using the Timed Up & Go test in a clinical setting to predict falling in Parkinson's disease. *Arch Phys Med Rehabil* 2013; 94: 1300–1305.
 33. Foreman KB, Addison O, Kim HS, Dibble LE. Testing balance and fall risk in persons with Parkinson disease, an argument for ecologically valid testing. *Parkinsonism Relat Disord* 2011; 17: 166–171.
 34. Huang SL, Hsieh CL, Wu RM, Tai CH, Lin CH, Lu WS. Minimal detectable change of the Timed "Up & Go" test and the dynamic gait index in people with Parkinson disease. *Phys Ther* 2011; 91: 114–121.