

Reliability and validity of floor transfer test in subjects with idiopathic Parkinson's disease

¹Abdul Basit, ¹Majumi M. Noohu, ²Fayaz R. Khan, ²Mohamed Faisal Chevidikunnan

¹Centre for Physiotherapy and Rehabilitation Sciences, Jamia Millia Islamia, New Delhi, India;

²Department of Physical Therapy, Faculty of Medical Rehabilitation Sciences, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia

Abstract

Objective: To determine the validity and reliability of the floor transfer test (FTT) in people with Parkinson's disease (PD). **Methods:** The sample was 25 subjects with PD. The concurrent and convergent validity of FTT was determined by correlation with scores of timed up and go test (TUG), movement disorder society- unified Parkinson's disability rating scale-III [motor component] (MDS-UPDRS), Schwab and England activities of daily living scale (SEADL). The intra rater reliability (ICC 3, 1), SEM, MDC, Bland & Altman limit of agreement (LOA) were determined. **Results:** FTT scores were positively correlated with TUG ($\rho=0.59$, $p=0.02$), MDS UPDRS III ($\rho=0.69$, $p=0.001$) and negatively correlated with SEADL ($\rho=-0.79$, $p=0.001$) showing moderate concurrent and convergent validity. ICC (3,1) was 0.92, SEM and MDC were 1.28 and 3.54 respectively.

Conclusion: FTT is a valid and reliable tool to assess functional mobility in people with mild PD.

Keywords: Mobility; balance; falls; movement disorder

INTRODUCTION

People with impaired balance is at higher risk of reduced physical activity, decline in functional independence and falls.¹ Balance impairment along with involvement of righting reactions, is a hallmark of idiopathic Parkinson's disease (PD).²⁻⁴ The commonly used scales to assess balance and functional mobility in people with PD are Berg balance scale (BBS), timed up and go test (TUG)⁵⁻⁷, Fullerton advanced balance scale (FABS)⁸ and balance evaluation systems test (BEST).⁹ Though these tests are multicomponent in nature neither of the tests have a component to assess the floor transfer ability.

Floor transfer test (FTT) is relatively a new clinical test of functional ability. It is a quick and easy to administer test which can be used in any environment and with minimal cost without special equipment.^{10,11} Either time taken by the subject to complete the task or the task is graded according to the ability of the subject's performance in the test.^{10,12} It was found to be a reliable and valid tool to assess functional mobility in subjects with stroke and community dwelling older adults. FTT is a good predictor of falls in community dwelling older adults as well.^{10,11,12}

There is dearth in availability of validated

tools for assessing the ability of people with PD, in changing the position from standing to sitting on the floor and then getting up back to standing. Hence this study investigated intra-rater test-retest reliability along with concurrent and convergent validity of FTT in people with PD. TUG, Movement Disorder Society-united Parkinson's disability rating scale (MDS-UPDRS)¹³ and Schwab and England activities of daily living scale (SEADL)¹⁴ tests were found to be appropriate to assess functional mobility, impairments and ADL respectively. TUG has established validity and reliability for assessing functional mobility in people with PD.^{15,16} Impairment can impact mobility, further ADL can be affected by mobility issues. So in this study TUG, MDS-UPDRS and SAEADL test, were used as reference to determine the validity.

METHODS

Subjects

Twenty five subjects diagnosed with PD were recruited from the department of physiotherapy and health center of Jamia Millia Islamia, New Delhi. Participants included in the study were both male and female over age of 40, Hoehn

& Yahr scale (H &Y) grade up to 2, duration of disease 1 year or longer, able to stand for at least 1 minute without support and able to ambulate with or without an assistive device for 10 meters. Individuals with atypical Parkinsonism, any previous surgical management of PD and other neurological co morbidities were excluded from the study. Proper information about the study purpose and associated risks with study had been explained to each participant, before the commencement of protocol. Subject's characteristics collected include age, gender, height (m), weight (kg) and medications. Two assessment tools, movement disorder society-unified Parkinson's disability rating scale-III [motor component] (MDS-UPDRS), and Schwab and England activities of daily living scale (SEADL) were used to check the baseline data for the disease severity and functional independence of the participant. If the subject was taking medications for Parkinson's disease, they were tested in the "on time" of medications. All subjects had signed an informed consent before participation in the study. This study was approved by the institutional ethics committee of Jamia Millia Islamia, New Delhi.

Sample size

A sample size of 25 subjects with 2 observations per subject achieves 86% statistical power to detect an intra-class correlation (ICC) of 0.70 under the alternate hypothesis and ICC under the null hypothesis of 0.30 with a significance level of 0.05, was required for reliability testing.¹⁷ A sample size of 19 subjects achieves 80% statistical power to detect a difference of -0.60 between null hypothesis correlation of 0.00 and alternate hypothesis of 0.06 with a significance level of 0.05, was required for validity testing.¹⁸

Procedure

Each participant had three assessment sessions. The subject with PD were evaluated for participation eligibility on day one. The demographic information and other disease details were collected from eligible subjects who gave the consent to participate in the study. FTT and TUG testing along with MDS-UPDRS and SAEDL evaluation was done on 2nd (measurement 1) and 7th day (measurement 2) of the initial assessment. Each assessment session lasted 45 to 60 minutes. The study protocol is given in Figure 1.

Prior to actual testing a practice trial of FTT and TUG was performed by the subject. Two

trials were recorded for both the tests. Time taken to complete the task in the test was measured in seconds. The best of recorded trial of the tests were used for final analysis. Three to five minutes of rest between each trial and ten minutes of rest were given between two tests to minimize fatigue. The order of TUG and FTT testing were determined by randomization of cards. Testing was stopped if the subject reported any problem in performing the tests. All the assessment and testing was done by a physiotherapist who had more than 5 years of experience in managing people with neuromuscular disorders. He was not aware of the purpose of the study.

Instruments

Floor transfer test: During FTT testing, the subject was made to stand and told to change position from standing to sitting on a floor mat, then return to the standing position. During the test the subject was allowed to adopt his own strategy for position change¹². The time taken in seconds for the entire task was recorded as FTT score, using a stop watch. A chair was placed nearby the subject for support if needed and an assistant was there with the participant to prevent falls during testing.

Timed up and go test: TUG test is a functional mobility assessment tool for older adults.¹⁹ It has established validity for assessing functional mobility in people with PD.¹⁵ TUG has high test retest reliability and intra-rater reliability in people with PD. The test can detect the difference in functional performance in people with and without PD.¹⁶ The TUG testing consisted of the subject getting up from a chair with armrest, walk a 3m pre marked distance on the floor in a straight line, turn around, walk back to the chair at self-selected speed and sit down. The subject had to get up from the chair from on the command "GO" by the examiner and finish the test as per the test protocol. The total time taken for entire activity was recorded in seconds as TUG score, using a stop watch.

Hoehn and Yahr Scale: H & Y scale is a commonly used clinical rating scale to assess motor function in people with PD. It is a simple and easy tool to captures the typical pattern of progressive motor impairments.²⁰ The scale categorizes the disease into one of five stages of disease progression. It is used as a gold standard for testing of newly developed scales, and has good correlation with UPDRS total score.²¹

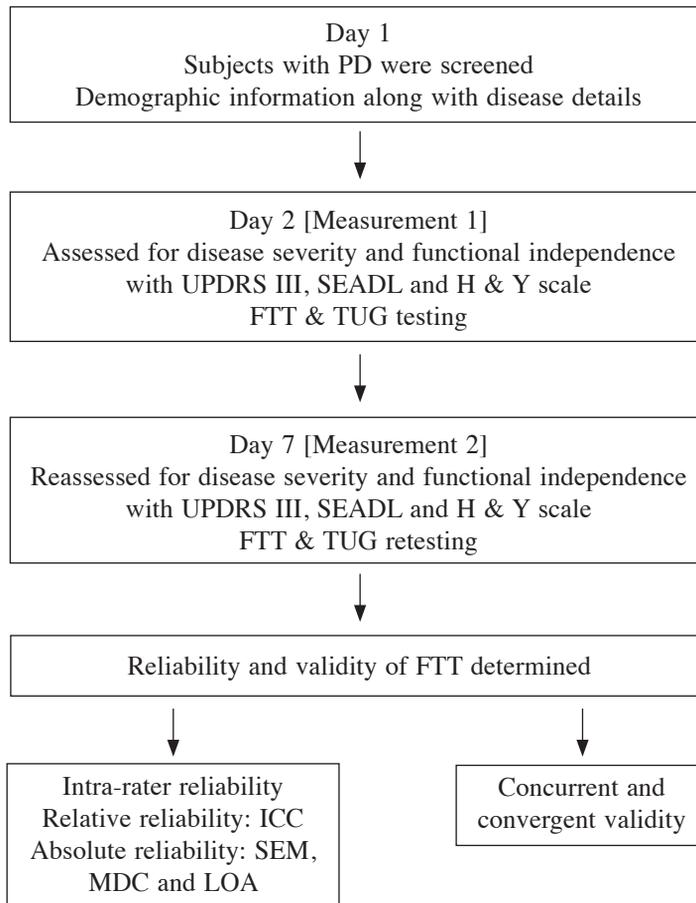


Figure 1. Protocol of the study

FTT - Floor Transfer Test; TUG -Timed up and go test; UPDRS III - Unified Parkinson's Disease Rating Scale III; SEADL - Schwab England activity of daily living scale, H & Y - Hoehn and Yahr scale; ICC- intraclass correlation; SEM- standard error of measurement, MDC- minimum detectable change, LOA- limits of agreement

MDS – Unified Parkinson's disease rating scale III: The UPDRS was revised in 2007 by the movement disorder society. The scale is as good as original version and included clinically pertinent questions making it more sensitive than the earlier version. The clinometric analysis supports the reliability and validity of the MDS-UPDRS. The internal consistency of part III motor examination was excellent ($\alpha= 0.93$) and it has excellent concurrent validity ($r=0.96$) with original UPDRS.²² The MDS-UPDRS III [motor component], consist of 18 questions related to motor functions. The scale was administered by the examiner and response was recorded as per the examination finding. Each question had five possible responses that are associated to frequently accepted clinical terms, 0=normal, 1=slight, 2=mild, 3= moderate, and 4= severe.

Schwab and England activities of daily living scale: The SEADL was initially used to assess response to basal ganglia surgery.²² It has adequate test retest reliability (ICC= 0.7) for use in people with PD. It has a standard error of measurement (SEM) of 4.45 and MDC₉₅ (minimal detectable change) of 12.33 in Parkinson's disease population^{14,22}, and it has adequate inter rater reliability (ICC=0.6) between physicians, patients and caregivers. The scale was administered by the evaluator and it measured the functional capacity, on a scale ranging from 0 to 100. Zero indicating worst possible function and 100 indicating no impairment.

Data analysis

The data were analyzed using SPSS 21 (SPSS Inc., Chicago, IL, USA). The descriptive statistics

Table 1: Characteristics of the participants. All the values are in Mean (SD) otherwise mentioned

Variables	Mean (SD)	Mean(SD)
	Day 2	Day 7
Age (yrs)	58.64(10.03)	-----
Duration of the disease (yrs)	2.94(1.85)	-----
Sex (Male: Female)	17:8	17:8
BMI (kg/m ²)	23.46(2.79)	-----
H & Y scale	1.76(0.43)	-----
SEADL	78.40 (7.46)	78.4(6.87)
MDS UPDRS III	24.20 (10.74)	24.48 (9.76)
TUG (s)	15.06 (3.53)	-----
FTT (s)	11.35 (4.53)	11.65 (5.62)
Medications*	1 (0)	-----
Type of medications (n)	25 (L)/ 5(ACIgc)	-----

SD= Standard deviation; SEADL: Schwab England Activities of daily living Scale; MDS UPDRS III: Unified Parkinson's Disease Rating Scale (Motor); H & Y: Hoehn & Yahr scale; TUG: Timed up and Go test; FTT: Floor transfer test , L: Levodopa, ACIgc: Anticholinergics, *- Median(Interquartile range: IQR)

(mean \pm SD/ n) of all the participants such as age (years), duration of disease (years), gender, BMI (kg/m²), H & Y scale, SEADL, MDS-UPDRS III [motor component], TUG (sec) and FTT (sec) scores were determined. The distribution of the data was determined and checked for normal distribution. Normal distribution of data was assessed using stem and leaf plot and Shapiro Wilk (S-W) test. The assumption of normality was not met for the FTT and TUG. The concurrent and convergent validity estimation were based on non-parametric tests using Spearman correlation analysis. The correlation of FTT scores with age, stage of the disease and the duration was also done using spearman correlation test. The FTT scores were log transformed for estimation of ICC as part of reliability testing. SEM, MDC and limits of agreement (LOA) were other parameter estimates of reliability testing. The subject's 2nd and 7th day MDS-UPDRS III and SEADL scores were compared using Wilcoxon sign ranked test. Significance level of $p \leq 0.05$ was set for all the analysis.

Validity analysis

Spearman correlation coefficient (ρ) was

calculated to measure the validity. Concurrent validity was determined by relating the scores on the tool of interest, with the score recorded on a standard measurement tool, which accurately measure that same construct. Convergent validity was estimated by comparing the scores on the test of interest with scores recorded with another tool, which quantifies a related but dissimilar construct. The FTT and TUG correlation was checked for concurrent validity, while the convergent validity was assessed using FTT and MDS-UPDRS III, FTT and SEADL. Correlation less than 0.5 were considered weak to fair, 0.5 to 0.75 moderate and greater than 0.75 were considered strong.

Reliability analysis

Intrarater test- retest reliability was analyzed by assessing relative and absolute reliability. The relative reliability was assessed by checking for systematic error, Spearman correlation coefficient and intra class correlation coefficient (ICC). The absolute reliability was checked by calculating standard error of measurement (SEM), minimum detectable change (MDC) and limits of agreement plots.

For non-normally distributed data, Wilcoxon

Table 2: Correlation of FTT with other tests

	Spearman correlation coefficient (ρ)	p
TUG	0.59	0.002
UPDRS III	0.69	0.001
SEADL	-0.79	0.001

Table 3: Summary of the statistical analysis for measurement 1 and 2 for measurement of Floor transfer test for intrarater reliability

Test	ICC (3, 1)	95% CI for ICC	SEM	MDC
FTT	0.92	0.83-0.96	1.28	3.54

ICC (3, 1) = Intra class correlation coefficient 2-way mixed single measure; CI= Confidence Interval; SEM= Standard Error of Measurement; MDC= Minimum Detectable Change; FTT= Floor Transfer Test

signed test was performed to check for any systematic error, by comparing FTT score on measurement 1 (day 2) and measurement 2 (day 7). It is assumed that there is systematic error in data if there was significant difference observed in Wilcoxon signed rank test. ICC 3, 1 (2-way mixed effect and consistency) for measurement 1 and 2 were computed, to determine the intrarater reliability. Intrarater reliability was considered to be acceptable for ICC >0.75 and considered to be very good for ICC >0.9.

The SEM was calculated as $SEM = SD \times \sqrt{1 - ICC}$; where SD is the standard deviation. A high level of SEM indicates a high level of error and implies non reproducibility of measurement. The minimum detectable change at 95% confidence interval was calculated to provide clinical interpretation as follows: $MDC = SEM \times 1.96 \times \sqrt{2} = 2.77 \times SEM$. Limit of agreement (LOA) between measurement 1 and 2 was calculated according to the procedure described by Bland and Altman.²⁴ LOA were expressed together with the mean difference between measurement 1 and 2, and

were judged whether they were narrow enough for the test to be of practical use.

RESULTS

The demographic characteristics of all the participants such as age, duration of disease, gender, details of medications BMI(kg/m²), H & Y scale, SEADL, MDS-UPDRS III [motor component], TUG and FTT scores is given in the Table 1. All subjects (n=25) were available for both the assessment 1 and 2. There was no correlation between the age and FTT score ($\rho=0.25$, $p=0.22$) and duration of the disease ($r=0.19$, $p=0.35$). The stage of the disease (HY stage) showed a moderate positive correlation with FTT scores ($\rho=0.47$, $p=0.01$). The reliability and validity estimation was based on data of all the subjects participated in the study. Shapiro Wilk test indicated that TUG, FTT and SEADL were not normally distributed whereas MDS-UPDRS III [motor component] score was normally distributed.

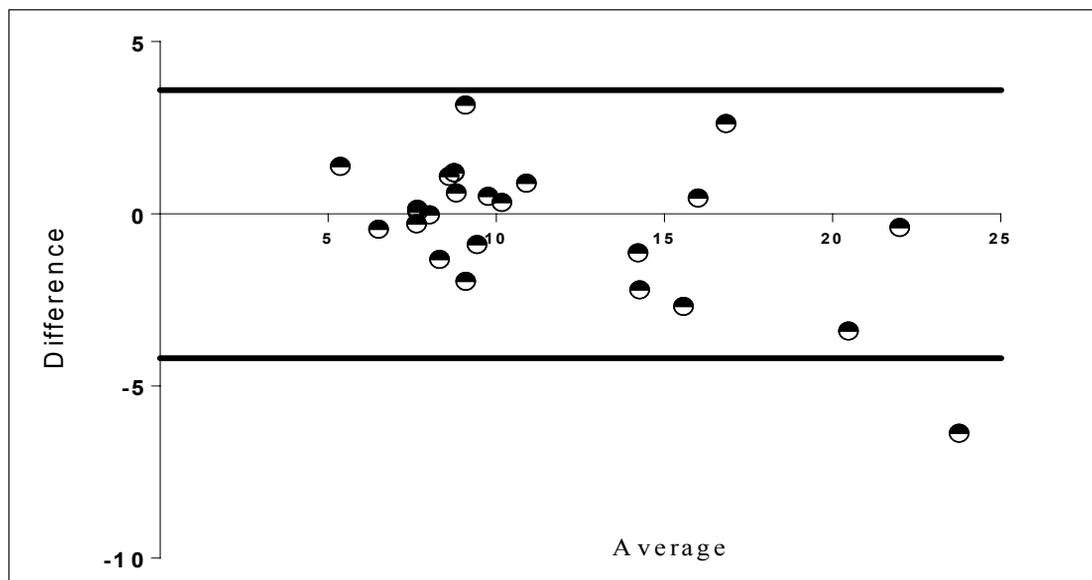


Figure 2. The Bland-Altman of agreement between test and retest of the FTT scores. In the plot except one point, all the data points were within ± 2 SD from the average, so there is 96% of agreement between the two sessions of measurement

Validity

The concurrent validity analysis of FTT score with TUG score showed a significant positive moderate correlation. The convergent validity analysis of MDS-UPDRS III [motor component] showed a significant positive moderate correlation, while SAEDL showed a significant negative strong correlation with FTT score. The correlation values between the tests scores is given in Table 2.

Reliability

The analysis using Wilcoxon signed rank test indicated that 7th day test score were statistically non-significant than 1st day score ($Z=-0.21$, $p=0.83$). The Spearman correlation coefficient for FTT was statistically significant ($r_s=0.86$, $p=0.001$) showing a good correlation with time scores obtained by the same rater on consecutive visit within the interval of 6 days. The intrarater reliability results together with 95% CI, for ICC, MDC and SEM results of the FTT scores are presented in Table 3.

The Bland-Altman (B&A) analysis plot between test and retest of the variables is shown in Figure 2. The mean difference determined by LOA analysis was very small. In the plot except one point, all the data points were within ± 2 SD from the average, so there is 96% of agreement between the two sessions of measurement. B&A plot is the quantification of the agreement between the two measurements by plotting it graphically, the mean difference and constructing limits of agreement. The limit of agreement recommended by B & A is that 95% of the data points should lie within ± 2 SD of the mean difference.²³

DISCUSSION

Assessment of functional mobility status at a particular point of time and continuous monitoring for the changes over the time in people with PD, helps in understanding the course of disease, patient's response to medication and rehabilitation interventions. The tools and scales with proven psychometric properties should be used for assessment, monitoring and treatment planning. Validity and reliability are the two components of psychometric property of any scale or tool. New assessment techniques should be checked for its validity and reliability in specific population. In this study we investigated the concurrent and convergent validity along with reliability of FTT in people with PD to assess floor transfer ability.

Concurrent validity is a type of criterion

validity and it tests how the scores from a new tool matches with scores of an established tool measuring identical construct. TUG being a standard tool for assessing functional mobility was selected as reference in this study. The result of the study showed that FTT has moderate concurrent validity with TUG in people with PD. There was a positive correlation between the scores of FTT and TUG scores, implying these two tests specifically evaluate functional mobility. Both FTT & TUG are similar tests in terms of assessing functional mobility related to patient's activities of daily living⁷. The moderate relationship between the FTT and TUG is similar to other studies that have assessed the validity of FTT in different population. FTT score was found to have a moderate correlation with functional reach distance, 50 ft. walk test and scores on performance oriented mobility test for balance.¹² In people with stroke it has shown a moderate correlation (0.75) with TUG scores.¹⁰

The convergent validity analysis showed a moderate correlation of FTT with assessment tools like MDS-UPDRS and SAEDL used in PD. Convergent validity is a type of a construct validity, its checked by comparing the scores obtained on the test with another test which measures a different construct. The MDS-UPDRS showed a positive correlation, whereas SAEDL scale had a negative correlation with FTT. The findings suggest that floor transfer ability can be influenced by motor symptoms and it can also pick up the level of activities of daily living. In one of the study, TUG compared with UPDRS III and H & Y scale showed moderate correlation¹⁵, which is also similar to our findings of relation of FTT with MDS-UPDRS and SEADL scale. The validation study of FTT in older adults have shown positive strong correlations between floor transfer ability and short physical performance battery scores ($r=0.87$) and physical functioning subscale ($r=0.87$). It was also observed in that study, older adults who were without physical disabilities and functionally independent, were able to pass FTT and those with physical disabilities and functionally dependent were not able to perform the test. It is useful to identify older adults with physical disabilities.¹¹

The FTT showed an excellent intra rater test-retest reliability (ICC). Our findings of the results are similar to previous works done by Murphy *et al.* and Ardali *et al.*, in older adults (ICC=0.79).^{5,11} Another study by Ng *et al.*, in patients with stroke found that FTT has excellent test retest reliability (ICC= 0.95).¹⁰ We found the test was highly

reliable, participants mean score was 11.35s on measurement 1 and 11.65s measurement 2. MDC calculated was 3.54s, to make 95% sure that the participants improved on FTT as an outcome measure, a minimum reduction of 3.54s should be present from the baseline measurement. A low value of 1.28 of SEM indicates that estimation made was steady and reproducible over time. As expected there was no change in disease severity and functional independence in the study participants over the period of 1 week.

It was observed that age and duration of the disease is not correlated with floor transfer ability. The stage of the disease showed a moderate positive correlation with FTT scores. The findings suggest that the age and duration of disease may not affect the mobility but the stage of disease can impact the mobility. There is involvement of more parts of body and also involvement of righting reflexes in higher stages of disease as per the H & Y scale.

Floor transfer is a complex motor activity requiring the body parts to be controlled more precisely to perform the task. In PD one of the major problems is inability in initiating and controlling the movements. The findings of the study reveals that PD affect the floor transfer task. The slowness in movement and changes in anticipatory and reactive postural control mechanism might be the contributing factors. People among various societies around the world, especially among Asian countries, perform daily living activities sitting on the floor. FTT makes an ideal tool for meaningful and clinically oriented evaluation. The time taken to complete the test can be an objective measure of functional mobility. The continuous monitoring of FTT scores may also help in observing the disease progression and effectiveness of intervention strategies for mobility impairment in PD population.

TUG is a functional mobility test, which consists of getting up and sitting down on a chair, walking and turning. Whereas FTT focusses on a particular activity. FTT is simple to administer but it's more challenging than TUG as it requires more strength, flexibility, balance, coordination and eccentric control. It may be difficult for people with profound impairments in motor functions. Hence along the various test measures available to assess the functional mobility in PD, FTT becomes a distinct tool of assessment. The delay in ability to stand up from floor may help the clinician to identify and manage functional limitations.²⁴ People who are identified with difficulty in floor transfers can

be given anticipatory intervention strategies to prevent fall and increase independence.

There were some limitations of the present study. The test measured the time needed for task completion, it didn't take into consideration of movement components and associated compensatory strategies. The performance of the test depends on lower limb strength, vestibular balance, attention, etc. which were not examined in this study. The study used a convenience sample rather than more robust methods of sampling techniques. The study didn't correlate the scores with any of balance measurement tools. The mobility can depend on the balance ability and other physical problems. Hence future studies are recommended for correlating FTT score with established balance assessment tools like BBS, FABS and BEST in people with PD. The study also didn't take into consideration other possible physical problems associated with PD. The findings may be used with caution in any other associated problems. Further studies also recommended with dual tasking, accuracy of FTT in predicting falls and testing the subjects in off phase of the medications.

In conclusion, the floor transfer test scores has excellent intrarater reliability and moderate concurrent and convergent validity. These findings indicate that that floor transfer test can be used as a valid and reliable tool to assess functional mobility in people with mild PD.

ACKNOWLEDGEMENT

The authors acknowledge the support of the administration of Jamia Milia Islamia, New Delhi, where the study was carried out. There was no conflict of interest involved.

REFERENCES

1. Park JH, Kang YJ, Horak FB. What is wrong with balance in Parkinson's disease? *J Mov Disord* 2015; 8(3): 109-14.
2. Wielinski CL, Erickson-Davis C, Wichmann R, Walde-Douglas M, Parashos SA. Falls and injuries resulting from falls among patients with Parkinson's disease and other Parkinsonian syndromes. *Mov Disord* 2005; 20(4): 410-5.
3. Antonini A, Barone P, Marconi R, *et al.* The progression of non-motor symptoms in Parkinson's disease and their contribution to motor disability and quality of life. *J Neurol* 2012; 259 (12): 2621-31.
4. Wood BH, Bilclough JA, Bowron A, Walker RW. Incidence and prediction of falls in Parkinson's disease: a prospective multidisciplinary study. *J Neurol Neurosurg Psychiatry* 2002; 72(6): 721-5.
5. Berg K, Wood-Dauphinee S, Williams JI, Maki B.

- Measuring balance in the elderly: Validation of an instrument. *Can J Public Health* 1992; 83: S7-S11.
6. Steffen T, Seney M. Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-item short-form health survey, and the unified Parkinson disease rating scale in people with parkinsonism. *Phys Ther* 2008; 88(6):733-46.
 7. Barry E, Galvin R, Keogh C, Horgan F, Fahey T. Is the Timed Up and Go test a useful predictor of risk of falls in community dwelling older adults: a systematic review and meta-analysis. *BMC Geriatr* 2014; 14: 14.
 8. Schlenstedt C, Brombacher S, Hartwigsen G, Weisser B, Möller B, Deuschl G. Comparing the Fullerton Advanced Balance Scale with the Mini-BESTest and Berg Balance Scale to assess postural control in patients with Parkinson disease. *Arch Phys Med Rehabil* 2015; 96(2): 218-25.
 9. Krzysztoń K, Stolarski J, Kochanowski J. Evaluation of balance disorders in Parkinson's Disease using simple diagnostic tests—not so simple to choose. *Front Neurol* 2018; 9: 932.
 10. Ng SS, Fong SS, Chan CW, *et al.* Floor transfer test for assessing people with chronic stroke. *J Rehabil Med* 2015; 47(6): 489-94.
 11. Ardali G, Brody LT, States RA, Godwin EM. Reliability and validity of the floor transfer test as a measure of readiness for independent living among older adults. *J Geriatr Phys Ther* 2019; 42(3):136-47.
 12. Murphy MA, Olson SL, Protas EJ, Overby AR. Screening for falls in community-dwelling elderly. *J Aging Physical Activity* 2003;11(1):66-80.
 13. Goetz CG, Tilley BC, Shaftman SR, *et al.* Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov Disord* 2008; 23(15): 2129–70.
 14. Bello-Haas VD, Klassen L, Sheppard S, 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.
 15. Verheyden G, Kampshoff CS, Burnett ME, *et al.* Psychometric properties of 3 functional mobility tests for people with Parkinson disease. *Phys Ther* 2014; 94(2): 230-9.
 16. Morris S, Morris ME, Ianseck, R. Reliability of measurements obtained with the Timed “Up, & Go” Test in people with Parkinson disease. *Phys Ther* 2001; 81(2): 810-8.
 17. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med* 1998; 17(1): 101-10.
 18. Zar JH. *Biostatistical analysis* (2nd ed.). Prentice-Hall Inc., Englewood Cliffs, New Jersey 1984.
 19. Podsiadlo D, Richardson S. The timed “Up & Go”: a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc* 1991;39(2): 142-8.
 20. Hoehn MM, Yahr MD. Parkinsonism: Onset, progression, and mortality. *Neurology* 1998; 50(2): 318-54.
 21. Zhao YJ, Wee HL, Chan YH, *et al.* Progression of Parkinson's disease as evaluated by Hoehn and Yahr stage transition times. *Mov Disord* 2010; 25(6): 710-6.
 22. McRae C, Diem, G, Vo A, O'Brien C, Seeberger L. Reliability of measurements of patient health status: A comparison of physician, patient, and caregiver ratings. *Parkinsonism Relat Disord* 2002; 8(3): 187-92.
 23. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; 8(2):135-60.
 24. Bridgewater KJ, Sharpe MH. Trunk muscle performance in early Parkinson's disease. *Phys Ther* 1998; 78(6): 566-76.