Diagnostic Accuracy of the 3-Meter Backward Walk Test in Persons With Parkinson Disease

Valerie A. Carter, DPT, PT, NCS, GCS; Becky G. Farley, PhD, MS, PT; Kay Wing, DPT, PT, NCS; Hendrik “Dirk” de Heer, PhD, MPH; Tarang K. Jain, PhD, PT

Purpose: To test diagnostic accuracy of the 3-Meter Backward Walk (3MBW) test for falls among 175 persons with Parkinson disease (PwP).

Methods: Retrospective data analysis was performed on time for 3MBW test, forward walk tests (Timed Up and Go, TUG, cognitive, and 10-Meter Walk), and the 5 times Sit-to-Stand test. A receiver operating characteristic curve analysis was completed to determine fall cutoff scores, and the optimal area under the curve (AUC) with sensitivity/specificity values was defined.

Results: The 3MBW test at 4.2 seconds had the highest AUC (AUC = 0.699). Persons with Parkinson disease with advanced disease were slower for each test; the 3MBW was the only test that was significantly different across disease severity.

Conclusion: The 3MBW test had the highest overall accuracy for retrospective falls and should be considered as part of battery of tests among persons with Parkinson disease.

Key words: balance, fall risk, Parkinson, sensitivity, specificity, 3-Meter Backward Walk, Timed Up and Go

Approximately 50% of persons with Parkinson disease (PwP) who fall require medical care.1 Persons with Parkinson disease who fall have increased morbidity and mortality, caregiver burden, utilization and cost of health care, and have a reduced quality of life.2,3 Developing and optimizing therapeutic strategies to prevent falls require identifying PwP at risk for falling. It is critical to develop outcome tools that can effectively identify PwP who are at risk of falling earlier in the Parkinson disease (PD) continuum.

Over the years, many risk factors for falls have been identified such as age, levodopa treatment, and motor fluctuations.9 However, these risk factors are common with most PwP and may not contribute to predicting which subgroup of PwP is at greatest risk of falls. Recent research indicates that backward walking might be a more challenging measure that is better at indicating level of impairment in gait than forward walking and forward walking with dual task.10 Persons with Parkinson disease have an increased axial stiffness and more limited sense of proprioception and verticality, resulting in falls in backward and lateral directions.11-15

Recently, Carter et al14 introduced a novel clinical measure for fall risk using the 3-Meter Backward Walk (3MBW) test in a sample of community-dwelling older adults. The authors proposed that a low risk of falls score of 3.0 seconds and a high risk of falls score of 4.5 seconds were more predictive of retrospective falls in the elderly than other popular fall-predictive tests such as the Five Times Sit to Stand (5×STS) and the Timed Up and Go (TUG) and may be a valuable addition to the current battery of clinical tools.14 However, fall risk cutoff values for the 3MBW have not been determined in neurologically impaired populations such as PwP. In addition, no study has documented the extent to whether the 3MBW times increase with disease progression. Therefore, the following retrospective study aimed to

1. compare scores in time (seconds) on the 3MBW test with other existing measures of fall risk related to forward walking (TUG regular, TUG cognitive, and 10-Meter Walk) and the 5×STS tests between PwP who reported falls versus who did not;
2. compare 3MBW and forward walking tests across disease severity in PwP; and
3. define optimal cutoffs for diagnostic accuracy of 3MBW and forward walking tests as well as the 5×STS in PwP who reported falling.

METHODS

Design and participants
This study looked at data collected from an ongoing medical record data set of clients with PD who attended a PD-specific clinic (see the Table). The retrospective data presented here represent clients who received an initial
evaluation by PD-specialized physical therapists from July 2013 through July 2018 (n = 175) as part of a larger ongoing medical record review. Study procedures were reviewed and approved by the Northern Arizona University Institutional Review Board.

**Measures**

The following demographic characteristics were included for analysis: age, gender, years with PD, disease severity as defined by Hoehn and Yahr (HY) stage of PD, and fall history within the past 6 months (yes/no). In addition to the 3MBW, the following tests and measures commonly used to determine risk of falls in PwP were also included:

*Timed Up and Go regular (TUG regular)* aims to assess mobility, balance, walking ability, and fall risk in older adults and PwP. The test begins with individuals seated in an armchair with their back against the back of the chair. When the clinician says “go,” the person stands up, walks 3 m, turns around, walks back, and sits down again. This test has been well established with fall cutoff time for PwP ranging from 8 to 12 seconds.\(^{16,17}\)

*Timed Up and Go cognitive (TUG cognitive)* aims to assess walking mobility, balance, and fall risk under dual task conditions in older adults and PwP. Participants complete the same procedure as in the TUG regular but counted backward by threes from a randomly selected number between 20 and 100. This test has been well established with a fall cutoff time range of more than 15 seconds in the elderly with a history of falls and 21.5 seconds in PwP with a history of falls.\(^{18}\)

*Five Times Sit to Stand (5xSTS)* aims to assess lower extremity strength as well as fall risk.\(^{19}\) The test starts in seated position with arms folded across the chest and the individual is instructed to stand up and sit down 5 times as quickly as possible while reaching full standing position between repetitions. A fall cutoff score for a community-dwelling older adult is 12 seconds whereas a cutoff of more than 15 seconds is indicative of recurrent fall risk. For PwP, a fall risk cutoff score of 12 seconds for a single fall and 15 seconds for multiple falls has been indicated.\(^{20}\)

*10-Meter Walk (10MW)* aims to assess forward gait speed, which can predict morbidity and mortality as well as 8 to 12 seconds.

### TABLE Demographics and Outcome Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>TUG Regular (Mean ± SD)</th>
<th>TUG Cognitive (Mean ± SD)</th>
<th>5xSTS (Mean ± SD)</th>
<th>10MW (Mean ± SD)</th>
<th>3MBW (Mean ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male (N = 123)</td>
<td>7.8 ± 2.5 n = 112</td>
<td>10.1 ± 4.0 n = 109</td>
<td>12.7 ± 4.2 n = 68</td>
<td>7.0 ± 1.8 n = 120</td>
<td>4.5 ± 2.2 n = 116</td>
</tr>
<tr>
<td>Female (N = 52)</td>
<td>7.6 ± 2.4 n = 46</td>
<td>10.1 ± 4.3 n = 45</td>
<td>11.7 ± 4.3 n = 32</td>
<td>6.9 ± 1.8 n = 48</td>
<td>4.7 ± 2.0 n = 47</td>
</tr>
<tr>
<td><strong>Age (decade)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.00</td>
<td>5.9 ± 1.2 n = 13</td>
<td>6.8 ± 1.6 n = 13</td>
<td>9.7 ± 2.3 n = 9</td>
<td>6.0 ± 1.2 n = 13</td>
<td>3.3 ± 1.5 n = 13</td>
</tr>
<tr>
<td>6.00</td>
<td>6.8 ± 1.8 n = 42</td>
<td>9.4 ± 4.0 n = 42</td>
<td>11.5 ± 4.1 n = 25</td>
<td>6.7 ± 1.7 n = 45</td>
<td>4 ± 1.6 n = 44</td>
</tr>
<tr>
<td>7.00</td>
<td>8.0 ± 2.4 n = 77</td>
<td>10.3 ± 4.0 n = 75</td>
<td>12.2 ± 4 n = 47</td>
<td>7.1 ± 1.8 n = 83</td>
<td>4.6 ± 2.1 n = 80</td>
</tr>
<tr>
<td>8.00</td>
<td>9.3 ± 3 n = 26</td>
<td>12.3 ± 4.1 n = 24</td>
<td>15.3 ± 4.3 n = 19</td>
<td>7.5 ± 1.8 n = 27</td>
<td>6.1 ± 2.6 n = 26</td>
</tr>
<tr>
<td><strong>Hoehn and Yahr Stage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>6.7 ± 2.2 n = 60</td>
<td>8.3 ± 3.6 n = 60</td>
<td>10.6 ± 3.8 n = 42</td>
<td>6.2 ± 1.2 n = 60</td>
<td>3.4 ± 1.4 n = 60</td>
</tr>
<tr>
<td>2.00</td>
<td>7.6 ± 1.8 n = 66</td>
<td>10.1 ± 3.2 n = 66</td>
<td>12.8 ± 3.8 n = 38</td>
<td>6.7 ± 1.3 n = 67</td>
<td>4.6 ± 1.8 n = 67</td>
</tr>
<tr>
<td>3.00</td>
<td>9.0 ± 2.4 n = 22</td>
<td>13.5 ± 4.6 n = 21</td>
<td>14.9 ± 4.3 n = 16</td>
<td>7.9 ± 2.2 n = 26</td>
<td>5.9 ± 2.5 n = 25</td>
</tr>
<tr>
<td>4.00</td>
<td>11.7 ± 2.9 n = 10</td>
<td>14.3 ± 5.2 n = 7</td>
<td>16.8 ± 3.1 n = 7</td>
<td>9.4 ± 1.9 n = 15</td>
<td>7.5 ± 2.3 n = 11</td>
</tr>
</tbody>
</table>

Abbreviations: 5xSTS, Five Times Sit to Stand; 10MW, 10-Meter Walk; 3MBW, 3-Meter Backward Walk; TUG, Timed Up and Go.
as functional ability. The time to complete a 10-m distance (with 5 m on either side) was recorded in seconds. Participants completed two trials for each normal and fast 10MW tests, as 10MW tests administered at both comfortable and “fast and safe” walking speeds have demonstrated excellent test-retest reliability in PwP (intraclass correlation coefficient, ICC > 0.96). 21

3-Meter Backward Walk (3MBW) aims to assess backward walking mobility, balance, and fall risk. The test starts with participants standing straight facing backward on a marked line on a smooth surface such as tile or wood. They are then instructed to walk backward as quickly but safely as possible and stop after 3 m (marked on the floor). Participants were permitted to look behind themselves if they desired and the examiner walked with the participant to ensure safety. A fall cutoff range of 3.0 seconds (low risk of falls) and 4.5 seconds (high risk of falls) for the normal elderly community has been previously recommended. 14

Fall history is a strong predictor for future falls in the elderly. 22 The primary dependent variable for the current study was whether a participant reported falling (dichotomized yes or no) when asked to answer the question (“Have you fallen in the past 6 months?”). For this study, a fall was defined as an event that resulted in a person coming to rest unintentionally on the ground or other level. 23 24 For a fall to count toward the fall status categorization, the participant had to recall at least an approximate date and specific circumstances surrounding the fall.

Analyses
Data were analyzed using the SPSS software package, version 26 (SPSS, IBM Corp, Armonk, New York) and MedCalc 19.1.7 (MedCalc Software, Ostend, Belgium). For each variable, extreme outliers (more than 2.0 standard deviations away from the mean) as assessed by boxplot were excluded from the analysis. Demographic characteristics of participants were summarized with descriptive statistics, including mean ± SD and frequency distributions. The forward and backward walking times were compared for age (50-80 years) by decade (5-8 retrospectively), disease severity (HY stages 1-4), and reported falls using a 1-way analysis of variance. Post hoc tests were performed with Bonferroni corrections to account for multiple comparisons. Sensitivity was calculated as the proportion of PwP with a fall history who were correctly identified by a fall risk assessment tool as having fall risk. Specificity was calculated as the proportion of PwP who did not have a fall history and were correctly identified by a fall risk assessment tool as not having fall risk. Receiver operating characteristic (ROC) curves for different tests and measures were examined, and optimal cut points were determined from the shortest distance from the upper left of the ROC plot. Overall accuracy was assessed by defining the optimal area under the curve (AUC). In all statistical comparisons, \( P < .05 \) was used as the criterion for statistical significance.

RESULTS
Demographics and baseline values
A total of 175 PwP participated, most of whom were males (70.8%) in their 60s and 70s (75.3%). Although not all participants completed all outcome measures, at least 153 of 175 participants completed each of the functional tests performed in this study. Only 100 people completed the 5×STS as this test was added later in the process of data collection. The number of extreme outliers excluded from analysis for each variable (TUG regular, \( n = 15 \); TUG cognitive, \( n = 13 \); 5×STS, \( n = 12 \); 3MBW, \( n = 7 \); 10MW, \( n = 2 \)) were most representative of people with advanced PD (i.e., 8/10 reported falls and freezing of gait and represented nearly 39% HY stage 4). Overall, males and females had similar walking speeds on all the functional tests. For example, males averaged 4.5 ± 2.2 seconds on the 3MBW and females 4.7 ± 2.0 seconds; males averaged 7.8 ± 2.5 seconds on the TUG and 7.0 ± 1.8 on the 10MW and females averaged 7.6 ± 2.4 and 6.9 ± 1.8 seconds (Table).

Aim 1: Comparison of PwP who reported falling versus those who did not
People who reported falling in the past 6 months were compared with people who did not report falling. Age (\( P = .168 \)) and gender (\( P = .490 \)) were not significantly different between people who fell and people who did not fall. However, disease severity was significantly different between people who fell and people who did not fall (\( P = .000 \)). In total, 67 people reported falling in the past year and these data are summarized in Figure 1 for each measure. People who fell had a significantly slower 3MBW of 5.5 seconds on average, compared with 4.0 seconds for people who did not fall (\( P < .001 \)). People who fell also had a significantly slower TUG regular (8.6 vs 7.3 seconds, \( P = .001 \)), TUG cognitive (11.5 vs 9.3 seconds, \( P = .002 \)), 5×STS (13.8 vs 11.8 seconds, \( P = .025 \)), and 10MWT (7.6 vs 6.6 seconds, \( P < .001 \)) (see Figure 1).

Aim 2: Comparison of forward and backward walk tests across disease severity
We tested the difference in times during forward and backward walking across disease severity using 1-way analysis of variance. The test times were statistically significantly different for all forward and backward walking tests across disease severity (3MBW: \( F_{12,150} = 21.249; P < .001 \); TUG regular: \( F_{12,150} = 19.223; P < .001 \); TUG cognitive: \( F_{12,150} = 13.553; P < .001 \); 5×STS: \( F_{3,96} = 7.016; P < .001 \); 10MW: \( F_{3,164} = 21.732; P < .001 \)). The post hoc analysis showed that 3MBW was able to detect significant differences between stages 1, 2, 3, and 4, whereas forward walking tests and the 5×STS were able to detect significant differences between
stage 1 and stages 3 and 4 only. For example, people in stage 1 walked 6.7 ± 2.2 seconds on the TUG, which increased to 7.6 ± 1.8 (stage 2), 9.0 ± 2.4, and 11.7 ± 2.9 seconds in stage 4, or a 74% increase in time from stage 1 to stage 4. For the 3MBW, the times were 3.4 seconds (stage 1), 4.6 ± 1.8, 5.9 ± 2.5, and 7.5 ± 2.3 seconds, respectively, or a 120% increase in time from stage 1 to stage 4.

Aim 3: Define optimal cutoffs for diagnostic accuracy

Using ROC curves, optimal cutoffs were evaluated for the forward and backward walk tests and fall history (yes or no) (Figure 2). The diagnostic accuracy, as determined by the AUC of forward and backward walking tests for detection of falls, in the descending order was 3MBW: 0.699 (Sensitivity 71.4%/Specificity 64%), 5×STS: 0.674 (Sensitivity 64%/Specificity 69.1%), TUG cognitive: 0.666 (Sensitivity 64.2%/Specificity 61.7%), 10MWT: 0.656 (Sensitivity 32.8%/Specificity 92.7%), and TUG regular: 0.647 (Sensitivity 67.9%/Specificity 58.5%). The optimal cutoff points in terms of sensitivity and specificity for falls in PwP for different tests and measures were determined to be TUG: 7.3 seconds, TUG cognitive: 9.2 seconds, 5×STS: 12.4 seconds, 10MWT: 6.6 seconds, and 3MBW: 4.2 seconds.

DISCUSSION

In the context of prior research reporting that 3MBW is more sensitive at evaluating risk of falls in the healthy community-dwelling elderly, this study aimed to assess the ability of the 3MBW to identify PwP who reported fall in the previous 6 months and to determine an appropriate fall cutoff score for PwP. Clinically, sensitivity is an important measure to assess fall risk. High sensitivity is preferred because it corresponds to more true positives and fewer false negatives, whereas the opposite is true for low sensitivity. It is important that PwP who have fall risk be identified as such (true positives) and that no PwP who have fall risk be identified as not having fall risk (false negative) so that appropriate fall risk measures can be taken.

To our knowledge, this is the first study to identify the diagnostic accuracy of the 3MBW as a fall risk assessment tool in PwP. In this study, 3MBW cutoff of 4.2 seconds was identified to be the most optimal for identifying falls in PwP, having the highest overall AUC of 0.699. A cutoff at 4.2 seconds would correctly identify almost 71% of PwP who reported falling, while correctly capturing most of the PwP who did not. These were much higher than the 10MWT at 6.6 seconds, which missed almost 67% of the PwP who reported falling, although correctly capturing most of the PwP (95%) who did not fall. 5×STS and TUG cognitive were closest in overall accuracy to the 3MBW and not significantly different in overall AUC. This corresponds to possible proposed fall cutoffs times on the 3MBW test for PwP of 4.0 seconds for low risk and of 5.5 seconds for high risk of previous falls.
The 3MBW test is designed to give insight to the ability of a person to initiate and sustain walking in the backward direction. For the elderly and PwP, walking backward is more difficult than walking forward.\textsuperscript{11,12,22,25-27} Persons with Parkinson disease are most unstable during backward movement and have difficulty with initiating movement in the backward direction, take more steps to regain balance after a backward perturbation, and demonstrate more pronounced changes during backward walking thanagematched controls.\textsuperscript{11,12,20,27} Recently, Kwon et al\textsuperscript{10} showed backward walking variability in de novo PwP to be more variable than forward walking, as well as forward walking with dual task conditions. They determined that backward walking, as the earliest indicator of gait variability, may be strongly associated with the risk of falling and disease progression.\textsuperscript{10} This is consistent with the data in this study. Only the 3MBW test was able to detect differences across all HY stages (1-2-3-4) and the only test that was sensitive enough to detect a difference between early stages (HY1 and 2) when compared with the TUG regular, TUG cognitive, 5×STS, and 10MWT. Therefore, along with the capacity of fall predictability, the 3MBW test could be used as an additional functional outcome measure to detect changes in PwP disease severity.

Limitations and future directions
Limitations of this study include the retrospective falls assessment (fall history), which is not as strong as a prospective analysis. Although retrospective fall assessment studies are not uncommon, a prospective assessment would allow validation of the utility of 3MBW test for prediction of falls and its ability to detect disease severity in PwP. We included PwP with advanced PD in the analysis because this is representative of individuals who attend physical therapy for assessment and fall intervention programming. However, as our data showed, this poses unique problems in the interpretation of cutoffs because of outliers (>2 SD). While these extremes may be truly representative of the off medication periods and freezing of gait episodes experienced in daily life, the scores are hard to interpret. For example, 8 of 10 of our outliers experienced falls and freezing of gait (FOG) in real life, but during the data collection process, there is no standardized process for adapting the data collection process (ie, how long do you wait for them to move before ending the trial; what is the amount of cueing or assistance needed to complete the trial). While they may all be high fall risks, future studies need to determine what measures (and adaptations) are most meaningful for identifying change in those individuals with advance PD who have the most potential to improve.

Second, the incidence of falls was self-reported and some participants’ recall of falls may be inaccurate. However, in this study, a 6-month period, instead of a 12-month period, was used to assess history of falls and may have helped with recall.\textsuperscript{20} Future studies are needed to compare the accuracy of recall at different time points and to develop a more standardized set of questions for eliciting recall that controls for cueing or bias.

Third, the categorical distinction between people who fell only once may be perceived as a limitation. While a history of multiple falls to assess tool validity is supported in the literature,\textsuperscript{16,29} it is not clear whether a single isolated fall over a 6-month period would justify fall risk designation, whereas multiple falls would. Fourth, the data were collected from 1 PD-specialized clinic where all physical therapists were PD specialists. This may limit generalizability of findings to other settings where therapists may not be as familiar with assessments in PwP. Additional research that uses prospective methods with a larger and more heterogeneous sample of PwP in a variety of settings is needed to validate these findings and improve generalizability. It would also be interesting to assess the 3MBW test for predicting falls in other patient populations (such as Stroke and Multiple Sclerosis). This would help determine whether fall risk could be better identified with the addition of the 3MBW to the battery of existing fall risk-screening tools.

Finally, the fall risk assessment measures of this study assessed only functional mobility and, therefore, do not fully assess the array of possible reasons for increased fall risk; it is well known that falling is a complex and multifaceted issue.

CONCLUSION
The study was the first to evaluate the diagnostic accuracy of the 3MBW test with fall history and compare its performance with existing measures of fall risk in PwP. Backward walking has been shown to be more difficult than forward walking, even the early stages of PD, and could be used as an additional functional outcome measure to detect changes in the disease severity of PwP. The 3MBW with a fall cutoff score of 4.2 seconds was able to predict falls similarly or more accurately than other clinical measures in PwP who reported a fall within the past 6 months. The 3MBW test should be considered in the battery of tests that is completed to predict falls in PwP.

References


