Dual Tasking With the Timed “Up & Go” Test Improves Detection of Risk of Falls in People With Parkinson Disease

Roisin C. Vance, Dan G. Healy, Rose Galvin, Helen P. French

Background. Falls are a common and disabling feature of Parkinson disease (PD). Early identification of patients at greatest risk of falling is a key goal of physical therapy assessment. The Timed “Up & Go” Test (TUG), a frequently used mobility assessment tool, has moderate sensitivity and specificity for identifying fall risk.

Objective. The study objective was to investigate whether adding a task (cognitive or manual) to the TUG (TUG-cognitive or TUG-manual, respectively) increases the utility of the test for identifying fall risk in people with PD.

Design. This was a retrospective cohort study of people with PD (N=36).

Methods. Participants were compared on the basis of self-reported fall exposure in the preceding 6 months (those who had experienced falls [“fallers”] versus those who had not [“nonfallers”]). The time taken to complete the TUG, TUG-cognitive, and TUG-manual was measured for both groups. Between-group differences were calculated with the Mann-Whitney U test. The discriminative performance of the test at various cutoff values was examined, and estimates of sensitivity and specificity were based on receiver operating characteristic curve plots.

Results. Fallers took significantly longer than nonfallers (n=19) to complete the TUG under all 3 conditions. The TUG-cognitive showed optimal discriminative performance (receiver operating characteristic area under the curve=0.82; 95% confidence interval [CI]=0.64, 0.92) at a cutoff of 14.7 seconds. The TUG-cognitive was more likely to correctly classify participants with a low risk of falling (positive likelihood ratio=2.9) (<14.7 seconds) and had higher estimates of sensitivity (0.76; 95% CI=0.52, 0.90) than of specificity (0.73; 95% CI=0.51, 0.88) at this threshold (negative likelihood ratio=0.32).

Limitations. Retrospective classification of fallers and nonfallers was used.

Conclusions. The addition of a cognitive task to the TUG enhanced the identification of fall risk in people with PD. The TUG-cognitive should be considered a component of a multifaceted fall risk assessment in people with PD.
Falls among patients with Parkinson disease (PD) lead to progressive loss of mobility, independence, and confidence that ultimately may reduce survival. Prevalence rates for falls in these patients greatly exceed those in age-matched controls and were reported to be as high as 87% in a longitudinal study of PD. Given the high prevalence and the negative impact of falls, it is critical to continue improving methods for accurately identifying people with the greatest risk of falls. The multifactorial nature of falls in people with PD means that the use of individual measures of balance or gait may be insufficient to accurately identify the risk of falls. Therefore, a combination of clinical and physiological measures is recommended to allow a comprehensive evaluation. The most useful outcome measures to include in an assessment of falls in people with PD remain unclear, but components such as a history of falling, disease severity, and mobility or balance measures have been recommended.

People with PD have been shown to be significantly more susceptible to gait changes when carrying out dual tasks than age-matched controls. These gait changes include reduced gait speed, reduced stride length, and stride pattern. Decreased automaticity and impaired attentional flexibility resulting from cortical and basal ganglion dysfunction may play a role in this gait impairment. Automaticity is a component of skilled movement that is used to evaluate whether a skill is performed with little demand on attentional resources. Higher levels of activity in premotor and prefrontal cortical areas of the brain, measured with functional magnetic resonance imaging, during the performance of dual tasks have been reported for patients with PD than for age-matched controls who are healthy. The authors suggested that the patients’ limits of capacity may have been exceeded, resulting in a deterioration in the automaticity of task performance. Although an association between dual-task performance during gait and fall risk has yet to be established, previous literature suggested that people with PD prioritize concurrent tasks over postural tasks during dual tasking. This scenario has been shown to decrease safety and increase fall risk in this population.

The measurement of automaticity allows consideration of the effects of dual tasks on performance. An automaticity index, which involves expressing the performance of dual tasks as a percentage of the performance of a single task, allows measurement of the dual-task cost and is a method for standardizing the automaticity measurement across research studies. The Timed “Up & Go” Test (TUG) is a simple, quick, and widely used clinical tool for the assessment of lower limb function, mobility, and fall risk. It has been identified as a valid and reliable measure of mobility in people with PD and is recommended by both the American Geriatric Society and the British Geriatric Society as a component of a multifactorial fall risk assessment. Although the TUG has been reported to be a moderately sensitive predictor of falls in people with PD, when used alone it may not be sufficiently accurate to identify fall risk.

The TUG was first modified by Lundin-Olsson et al to determine the effect of a second task on balance and gait in older adults who were frail. A manual task of carrying a glass of water was combined with the TUG, creating a dual-task condition (TUG-manual). People with a difference of 4.5 seconds or more between the single task and the dual tasks were identified as being more frail and having a greater risk of falling. However, the diagnostic accuracy of this test was not examined in that study.

The TUG was further modified in a study of elderly people who had experienced falls (“fallers”) to include a cognitive task (TUG-cognitive) that involved counting backward in threes while completing the TUG. The study examined both types of dual-task TUG and investigated whether the TUG-manual and the TUG-cognitive were more sensitive and specific predictors of falls in elderly people who dwelled in the community than the TUG alone. The authors concluded that the TUG was sensitive (87%) and specific (87%) for identifying fallers and that the addition of neither a manual task nor a cognitive task increased the overall ability of the test to predict falls. Because people with PD are more susceptible to gait changes under dual-task conditions than age-matched elderly controls, it cannot be assumed that these findings can be applied to people with PD.

The TUG-manual and the TUG-cognitive are now widely used in clinical practice and have been included as components of larger outcome measures, such as the Modified Parkinson Assessment Scale and the Balance Evaluation Systems Test (BESTest). The benefits of adding an additional component to the TUG for fall risk assessment in people with PD have yet to be fully investigated.

The primary aim of this study was to investigate whether adding a task to the TUG increases the ability of the test to identify fall risk in people with PD. A secondary aim was to evaluate the optimal discriminative value of the test under single-task and dual-task conditions. The third aim was to ascertain whether the automaticity of movement during
To address these aims, we generated 3 hypotheses. First, adding a task to the TUG would increase the sensitivity and specificity of the test for identifying fall risk in people with PD. With regard to the second aim, it was hypothesized that fallers would be significantly slower than nonfallers at completing all 3 types of TUG. The third hypothesis was that the automaticity of gait during dual-task TUGs would be significantly greater in nonfallers with PD than in fallers with PD and that people would have greater difficulty automatizing their gait with a cognitive task than with a manual task.

**Method**

**Participants**

People with a diagnosis of PD were consecutively recruited from a tertiary hospital in Dublin, Ireland, between October 2011 and January 2012. All participants provided written consent and met the following inclusion criteria: clinical diagnosis of idiopathic PD; history of no falls or 2 or more falls in the preceding 6 months; ability to provide written consent; ability to be independently mobile for at least 9 m (30 ft), with or without a gait aid; Hoehn and Yahr stages 2 to 4; ability to provide 3 of 5 correct answers on the attention and calculation section of the Mini-Mental State Examination (MMSE); and history of falls or 2 or more falls in the preceding 6 months. A fall was defined as “an event that resulted in the participant unintentionally coming to the ground or other lower level.” The sample size calculation for this study was based on the difference in TUG speed between fallers with PD and nonfallers with PD because data on dual-task TUGs were not available. A mean difference in TUG speed of 5.6 seconds (SD=5.2) was used to calculate the sample size. It was determined that a sample of 36 was required to provide a statistically significant difference in TUG results between fallers and nonfallers (power of 90%; P≤.05).

**Procedure**

All assessments were conducted in a physical therapy gymnasium by the principal investigator (R.C.V.) and took 35 minutes to complete. So that the effect of motor fluctuation would be minimized, all assessments took place within 2 hours of participants taking antiparkinsonian medication. Demographic data were collected, and then cognitive, balance, and mobility assessments were performed. The MMSE was carried out to ensure that participants had sufficient cognitive function to complete the dual-task assessments.

The 3 types of TUG (TUG, TUG-cognitive, and TUG-manual) were timed with a stopwatch. The 3 tests were completed in random order, and each was repeated 3 times, with the average of the 3 scores being used for data analysis. The original TUG evaluated functional mobility by measuring the time taken to stand from a chair, walk 3 m, turn around, return to the chair, and sit down. The TUG-manual entailed completing the TUG while carrying a glass of water in one hand. The TUG-cognitive involved completing the TUG while counting backward in threes from a random start point. An automaticity index was calculated to assess the impact of both dual tasks on fallers and nonfallers. The speed under the dual-task condition was expressed as a percentage of the speed under the single-task condition: (single task/dual task) × 100. A maximal achievable score of 100% indicated no decrease in performance under the dual-task condition. This index was developed as a method for standardizing the assessment of dual-task performance and its effect on automaticity across research studies.

**Data Analysis**

Data were analyzed with STATA (version 12, Stata Corp, College Station, Texas). Descriptive statistics, including means and standard deviations, were computed when appropriate. Mean differences between fallers and nonfallers were tested for significance by use of independent t tests or the Mann-Whitney U test for continuous variables with parametric or nonparametric distributions, respectively. Various explanatory variables (age, sex, MMSE score, Berg Balance Scale score, scores on the 3 types of TUG, and Hoehn and Yahr stage) were analyzed by use of univariate logistic regression to identify potential factors contributing to the risk of falling, with “faller” or “nonfaller” being the outcome of interest. The independent variables identified as significantly associated with falls by univariate analysis (P<.05) were entered into stepwise multiple logistic regression analyses to determine the best explanatory independent variables.

The discriminative value of the TUG was examined by use of 2×2 tables to calculate estimates of sensitivity and specificity and associated positive and negative likelihood ratios. Receiver operating characteristic curve analyses with 95% confidence intervals (CIs) also were performed to describe model discrimination for all 3 types of TUG. The point that
simultaneously maximized sensitivity and specificity was selected as the cutoff value. Accuracy was calculated on the basis of the proportion of cases correctly classified with the cutoff value. The c statistic (area under the curve) ranged from 0.5 (no discrimination) to a theoretical maximum of 1. Values between 0.7 and 0.9 represented moderate accuracy, and values greater than 0.9 represented high accuracy. A c statistic of 1 represented perfect discrimination, indicating that scores for all fallers were higher than those for all nonfallers, with no overlap. Statistical significance was set at $P<.05$.

### Results

Thirty-six people (21 men and 15 women) were recruited. Demographic details are shown in Table 1. The average age of the participants was 71.4 years ($SD = 8.3$, range = 66 – 87). Participants in the faller group had an average of 3.65 falls in the preceding 6 months (range = 2–8 falls), had significantly higher Hoehn and Yahr stages, and were more dependent on mobility aids than those in the nonfaller group (Tab. 1). There were no significant between-group differences in the age ($P = .25$) or MMSE scores of the participants. With regard to the first hypothesis, the diagnostic accuracy statistics for the 3 tests are shown in Table 2. The TUG-cognitive had the highest sensitivity (76.5%) and specificity (73.7%) for identifying fall risk in participants with PD. The sensitivities of the TUG and the TUG-manual were lower, indicating that these tests alone were insufficient for identifying participants at risk of falls (Tab. 2). The highest likelihood ratio (2.9; 95% CI = 1.3, 6.5) was found for the TUG-cognitive. This value indicated that a faller was 2.9 times more likely to be categorized correctly with this test than a nonfaller (Tab. 2). The second

---

**Table 1.**

Baseline Characteristics of “Nonfallers” and “Fallers”

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Participants (N=36)</th>
<th>Nonfallers (n=19)</th>
<th>Fallers (n=17)</th>
<th>Mean Difference (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>71.4$^{ab}$</td>
<td>69.9 (6.56)$^{ab}$</td>
<td>73.12 (9.87)$^{ab}$</td>
<td>3.22 (−2.4, 8.9)</td>
<td>.25</td>
</tr>
<tr>
<td>ABC-6, %</td>
<td>52$^a$</td>
<td>58.8 (31.2)$^a$</td>
<td>44.4 (15.5)$^a$</td>
<td>−14.44 (−31.5, 2.5)</td>
<td>.09</td>
</tr>
<tr>
<td>BBS score (0–56)</td>
<td>50.4$^c$</td>
<td>52 (6)$^c$</td>
<td>51 (7)$^c$</td>
<td></td>
<td>.17</td>
</tr>
<tr>
<td>MMSE score (0–30)</td>
<td>27.2$^c$</td>
<td>28 (2)$^c$</td>
<td>27 (4)$^c$</td>
<td></td>
<td>.16</td>
</tr>
<tr>
<td>Men</td>
<td>21</td>
<td>9 (47.4)</td>
<td>12 (70.6)</td>
<td></td>
<td>.04$^d$</td>
</tr>
<tr>
<td>Hoehn and Yahr stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.009$^d$</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>10 (52.6)</td>
<td>1 (5.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>4 (21.1)</td>
<td>6 (35.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>5 (26.3)</td>
<td>10 (58.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking aid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.02$^d$</td>
</tr>
<tr>
<td>None</td>
<td>22</td>
<td>15 (78.9)</td>
<td>7 (41.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking stick</td>
<td>10</td>
<td>4 (21.1)</td>
<td>6 (35.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two walking sticks</td>
<td>2</td>
<td>0</td>
<td>2 (11.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking frame</td>
<td>2</td>
<td>0</td>
<td>2 (11.8)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Data are reported as number (percentage) of participants unless otherwise indicated. Nonfallers = participants who had not experienced falls, fallers = participants who had experienced falls, CI = confidence interval, ABC-6 = modified Activities-Specific Balance Confidence Scale, BBS = Berg Balance Scale, MMSE = Mini-Mental State Examination.

$^b$ Reported as mean (SD).

$^c$ Reported as median (interquartile range).

$^d$ Statistically significant at $P<.05$.

---

**Table 2.**

Sensitivity, Specificity, and Likelihood Ratio (LR) for 3 Types of Timed “Up & Go” Test (TUG)

<table>
<thead>
<tr>
<th>Variable (Cutoff Time, s)</th>
<th>AUC (95% CI)</th>
<th>% Sensitivity (95% CI)</th>
<th>% Specificity (95% CI)</th>
<th>Positive LR (95% CI)</th>
<th>Negative LR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TUG (12)</td>
<td>0.77 (0.60, 0.89)</td>
<td>41 (32, 51)</td>
<td>73 (64, 82)</td>
<td>1.57 (0.6, 4.1)</td>
<td>0.8 (0.5, 1.3)</td>
</tr>
<tr>
<td>TUG-cognitive (14.7)</td>
<td>0.82 (0.64, 0.92)</td>
<td>76.5 (52, 90.4)</td>
<td>73.7 (51.2, 88.2)</td>
<td>2.9 (1.3, 6.5)</td>
<td>0.32 (0.1, 0.8)</td>
</tr>
<tr>
<td>TUG-manual (13.2)</td>
<td>0.78 (0.61, 0.90)</td>
<td>29.55 (13, 53.1)</td>
<td>68.4 (46, 84.6)</td>
<td>0.9 (0.34, 2.5)</td>
<td>1.1 (0.67, 1.6)</td>
</tr>
</tbody>
</table>

$^a$AUC = area under the curve, CI = confidence interval, TUG-cognitive = TUG with an added cognitive task, TUG-manual = TUG with an added manual task.
hypothesis was confirmed by a Mann-Whitney U test comparing speeds of TUG completion by fallers and nonfallers. The results indicated that fallers were significantly slower at completing all 3 types of TUG (Tab. 3). Receiver operating characteristic curve analyses for the different forms of TUG (Figs. 1, 2, and 3) were used to determine the cutoff values that would maximize the sensitivity and specificity of each test. As shown in Table 2, the optimal cutoff times for discriminating fallers from nonfallers were 12 seconds, 14.7 seconds, and 13.2 seconds for the TUG, TUG-cognitive, and TUG-manual, respectively.

Calculation of the automaticity index for each dual-task TUG showed that the TUG-cognitive had significantly lower automaticity values ($P < .05$) for fallers (75.9%) than for nonfallers (83.7%), confirming the third hypothesis. The results indicated that participants with PD and a history of falls had the greatest difficulty automatizing their gait pattern when carrying out a cognitive dual task. There was no significant difference ($P = .7$) in automaticity between fallers (89.1%) and nonfallers (86.3%) for the TUG-manual.

Finally, univariate regression analyses demonstrated that sex and Hoehn and Yahr stage were the only baseline measures predictive of falls, with the odds of a fall being significantly higher in men than in women (odds ratio = 4.2; 95% CI = 1.02, 16.67). Similarly, progression in Hoehn and Yahr stages was associated with an increased likelihood of a fall (odds ratio = 3.6; 95% CI = 1.38, 9.26). When both variables were entered into a multiple regression model, only Hoehn and Yahr stage remained significant (odds ratio = 3.51; 95% CI = 1.28, 9.58).

Table 3. Group Differences in 3 Types of Timed “Up & Go” Test (TUG)$^a$

<table>
<thead>
<tr>
<th>Test</th>
<th>Median (Interquartile Range) for:</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nonfallers</td>
<td>Fallers</td>
</tr>
<tr>
<td>TUG</td>
<td>10.12 (6.6)</td>
<td>13.01 (10)</td>
</tr>
<tr>
<td>TUG-cognitive</td>
<td>12.09 (9.8)</td>
<td>17.14 (17.6)</td>
</tr>
<tr>
<td>TUG-manual</td>
<td>11.72 (6)</td>
<td>14.6 (13.3)</td>
</tr>
</tbody>
</table>

$^a$ Nonfallers = participants who had not experienced falls, fallers = participants who had experienced falls, TUG-cognitive = TUG with an added cognitive task, TUG-manual = TUG with an added manual task.

$^b$ Statistically significant at $P < .05$.

Figure 1. Receiver operating characteristic (ROC) curve for Timed “Up & Go” Test (original).

Discussion

The present study demonstrated that the addition of a cognitive task to the TUG enhanced the ability of the test to correctly identify participants with a high or low risk of falls. A cutoff of 14.7 seconds for the TUG-cognitive was found to optimize the discriminative performance of the test. The TUG and TUG-manual had lower prognostic accuracies than the TUG-cognitive for identifying participants at risk of falls. Furthermore, the results demonstrated that the automaticity of gait declined to a greater degree with the addition of a cognitive task rather than the addition of a manual task to the TUG.

The diagnostic accuracy of each test, including sensitivity and specificity, was calculated. Test sensitivity was a metric of how often the TUG detected fall risk for a participant in the faller group. Specificity, on the other hand, indicated how often the participant was not at risk of falling and was categorized as a nonfaller. Cutoff values were used to differentiate fallers from nonfallers and were calculated from receiver operating characteristic curves for each TUG. The original TUG had low sensitivity (41%) and higher specificity (73%) at a cutoff of 12 seconds. This result can be compared with those of 2 previous studies investigating fall risk factors in people with PD. Those studies demonstrated TUG sensitivity of 63% and specificity of 65% and TUG sensitivity of 65% and specificity of 53%. Both studies...
concluded that the TUG accuracy rate was not sufficient to predict fall risk.

The TUG-manual also had low sensitivity (29%) and moderate specificity (68%). The overall discriminative performance (accuracy) of the TUG-cognitive was confirmed by receiver operating characteristic curve analyses to be about 82%. Thus, the TUG-cognitive performed better than the other 2 types of TUG, with the highest sensitivity (76.5%) and specificity (73.7%) for identifying fall risk in participants with PD. Although the TUG-cognitive had moderate accuracy, it is unlikely to be sufficient as a sole test of fall risk. Sensitivity or specificity values that qualify as “high” are usually defined as values of 95% or greater. A combination of disease-specific, balance, and mobility measures is necessary to accurately identify fall risk in people with PD. The TUG-cognitive may be more useful as a component of a multifactorial fall risk assessment including other variables, such as history of falls, disease severity, and freezing of gait. These factors have been identified as having strong predictive qualities for fall risk in people with PD.

The differences identified in the discriminative abilities of the TUG-cognitive and the TUG-manual in the present study may have been due, in part, to the complexity of the task. The degree of difficulty of the secondary task has been reported to affect the pattern and speed of gait in people with PD. The secondary tasks chosen are widely used in clinical practice as standard additions in the dual-task TUG. The automaticity index indicated that the TUG-manual required less attentional demand than the TUG-cognitive for both fallers and nonfallers. This result may indicate that the secondary task was not sufficiently complex to reach the threshold attention needed to negatively affect the basic TUG. Overall, however, although differences in task complexity between the manual task and the cognitive task may have contributed to the discriminative abilities of the tests in the present study, we believe that our interpretation of the data is valid on the basis of the strength of the results reported. Future research comparing the complexities of TUG-manual secondary tasks are likely to add to the current findings. Furthermore, a comparison of automaticity values in fallers and nonfallers has yet to be done thoroughly in people with PD; therefore, further research on the use of automaticity as a fall risk measure is needed.

The results of the present study identified a significant difference in test
performance between fallers and nonfallers for all 3 types of TUG. These findings are in keeping with those of previous research on elderly people, indicating that fallers were slower than nonfallers at completing the TUG under single- and dual-task conditions. Nonfallers had lower Hoehn and Yahr stages than fallers, indicating lower disease severity. The relationship between disease severity and risk of falling was investigated in a recent meta-analysis of patients with PD. The findings suggested that the Hoehn and Yahr scale was a better predictor of falls than the Unified Parkinson Disease Rating Scale (UPDRS); however, neither was identified as a strong predictor of falls, potentially because of “the U-shaped relationship of PD with falls.” In this phenomenon, the rate of falls increases initially with disease severity, becomes stable as compensatory strategies are engaged, and eventually decreases further because of immobility. The fact that the participants in the present study were independently mobile, with or without an aid, may have resulted in the exclusion of people in the plateau stage or people who had become immobile.

Prospective recording of falls remains the reference standard for evaluating fall risk; therefore, retrospective categorization of fallers and nonfallers was a methodological limitation of the present study. Although the present study was a retrospective cohort study, efforts were made to minimize recall bias by limiting the time frame for inclusion and excluding people with a history of 1 fall. A history of 2 or more falls in the preceding 6 months was required for inclusion in the faller group. Future prospective cohort studies with larger sample sizes will allow for the exploration of more predictor variables and would further validate the current study findings. Additional research should compare the impacts of TUG dual tasks of various complexities on identifying fall risk.

In conclusion, the identification of fall risk is a critical step in developing effective therapeutic interventions to reduce falls in people with PD. The TUG-cognitive was shown to have better discriminative ability than the original TUG or the TUG-manual. The optimal cutoff times for discriminating fallers and nonfallers were 12 seconds for the TUG, 14.7 seconds for the TUG-cognitive, and 13.2 seconds for the TUG-manual. These findings support the use of the TUG-cognitive as a simple screening test with moderate sensitivity and specificity in a multifaceted fall risk assessment in people with PD. The low sensitivity and specificity of the TUG-manual indicated that their use in clinical practice should be limited. Finally, the automaticity value of the TUG in fallers was significantly lower than that in nonfallers and may have been a factor compounding fall risk.

Ms Vance, Dr French, and Dr Galvin provided concept/idea/research design. Ms Vance and Dr Galvin provided writing. Ms Vance provided data collection and analysis, project management, and facilities/equipment. Dr Healy provided participants. Dr French and Dr Galvin provided consultation (including review of manuscript before submission). Approval for the study was granted by the Beaumont Hospital Research Ethics Committee.


References


Dual Tasking With TUG in People With Parkinson Disease


