Predicting the Probability of Falls in Community-Residing Older Adults Using the 8-Foot Up-and-Go: A New Measure of Functional Mobility

Debra J. Rose, C. Jessie Jones, and Nicole Lucchese

The purpose of this study was to determine whether performance on the 8-ft up-and-go test (UG) could discriminate between older adult fallers ($n = 71$) and nonfallers ($n = 63$) and whether it would be as sensitive and specific a predictor of falls as the timed up-and-go test (TUG). Performance on the UG was significantly different between the recurrent faller and nonfaller groups ($p < .01$), as was performance on the TUG ($p < .001$). Older adults who required 8.5 s or longer to complete the UG were classified as fallers, with an overall prediction rate of 82%. The specificity of the test was 86% and the sensitivity was 78%. Conversely, the overall prediction rate for older adults who completed the TUG in 10 s or longer was 80%. The specificity of the TUG was 86% and the sensitivity was 71%.

Key Words: prediction, aging, physical activity

Much of the loss in functional mobility leading to increased disability and, eventually, falls is preventable or at less treatable through early detection and targeted interventions (Chaves, Garrett, & Fried, 2000; Gill, Williams, Richardson, & Tinetti, 1996; Hindmarsh & Estes, 1989; Lawrence & Jette, 1996; Lord, Clark, & Webster, 1991; Morey, Pieper, & Cornoni-Huntley, 1998). The assessment of dynamic balance and mobility is an essential component of any geriatric assessment of older adults, particularly those at risk for falls. Although predictors highly associated with falls are multifactorial and interactional in nature (Rubenstein & Robbins, 1984; Tinetti, Williams, & Mayewski, 1986; Woollacott, 2000), finding an easy, quick, and cost-efficient measure to identify potential fallers would help researchers and practitioners better identify individuals in need of a more comprehensive balance assessment.

The timed version of the up-and-go test (TUG; Podsiadlo & Richardson, 1991) originally developed by Mathias, Nayak, and Issacs (1986) serves as a simple and objective measure of dynamic balance and mobility. The test also assesses lower body strength and power, two additional components that contribute to good balance. The TUG requires a participant to stand up from a seated position, walk 3
m (9.84 ft) as quickly and as safely as possible to a line marked on the floor, turn around, walk back to the chair, and sit down in the shortest time possible. The test has content validity, in that it incorporates fundamental movements required for independent living, and good concurrent validity. It has previously been shown to correlate well with other tests measuring balance, gait, and more general physical activities associated with daily living (Podsiadlo & Richardson). In the original study conducted by Podsiadlo and Richardson, performance on the TUG could discriminate between functionally independent and dependent older adults. Participants who were able to complete the test in less than 20 s \((n = 17)\) were considered to be independent in transfer tasks associated with basic activities of daily living. They also scored in the upper third on the Berg Balance Scale (Berg, Wood-Dauphinee, Williams, & Maki, 1992), another measure commonly used to assess functional limitations in balance, and walked at a speed that was sufficient for safe community ambulation. In contrast, adults who required longer than 30 s \((n = 26)\) to complete the test tended to be much more functionally dependent. Most required assistance with basic transfers, only 1 participant in the group could climb a set of stairs independently, and none were able to go outside without requiring some type of assistance.

More recently, researchers have also demonstrated the TUG’s ability to accurately classify community-dwelling older adults as fallers or nonfallers (Gunter, White, Hayes, & Snow, 2000; Shumway-Cook, Brauer, & Woollacott, 2000). Gunter and colleagues found the TUG to be the best predictor of faller status when compared with other clinical tests of mobility and laboratory assessments of strength and power. The TUG demonstrated very high sensitivity, with 98% of 109 fallers being accurately classified, but considerably lower specificity, with only 31% of 48 nonfallers being correctly classified. In contrast, Shumway-Cook et al. were able to correctly classify 87% of the 15 recurrent fallers and 87% of the 15 nonfallers who participated in their study. Two notable differences between the two studies were that the sample size was considerably smaller and the inclusion criteria for fallers more stringent in the Shumway-Cook et al. study. In that study, the inclusion criteria for fallers was a history of two or more falls within the previous 6 months, whereas Gunter’s group included older adults who had fallen one or more times within the past year in their faller group. These two factors no doubt contributed to the significantly higher overall classification rate obtained (87% vs. 65.5%) in the Shumway-Cook et al. study.

Shumway-Cook et al. (2000) further demonstrated that older adults who required 13.5 s or longer to perform the TUG were classified as fallers, with an overall correct prediction rate of 90%. The cutoff value established in this study is considerably lower than the 30-s value that Podsiadlo and Richardson (1991) found was best for predicting functional dependence in older adults. The authors attributed the differences to the fact that the earlier study included a large number of older adults with known neurologic diseases.

Although the TUG is a widely used measure of functional mobility and appears to be a sensitive and specific predictor of community-dwelling older adults who are most likely to fall, Rikli and Jones (1999a) found that the distance required to perform the test was often reported to be too long for the test to be easily and safely administered in the home. In addition, many of the older adult participants in their early pilot studies found certain test instructions confusing, particularly as related
to the turnaround area at the end of the 3 m. For example, participants were unsure whether both feet had to cross the line or just one foot. The use of a line to mark the turnaround has also been found to promote similar confusion among older adults who perform the test using a walker.

In order to increase the feasibility of administering the test in a home setting and reduce the confusion regarding the turning area in the TUG, Rikli and Jones (1999a) reduced the overall distance required to perform the test to 8 ft (2.44 m) and replaced the line with a cone that signaled the turnaround point. The modified version of the TUG is referred to as the up-and-go test (UG). This shorter version of the test requires participants to rise from a chair, walk 8 ft to and around a cone, walk back, and sit down in the shortest time possible. Similar to the TUG, the test–retest reliability for the shorter version test (UG) has also been reported to be very high (.95; Rikli & Jones, 1999a).

A second advantage of the UG over the TUG is that it constitutes one test item within a more comprehensive functional-fitness test battery designed to assess multiple physical parameters associated with functional mobility. The complete test battery consists of five additional items (including one alternative test). Test items include the chair stand (a measure of lower body muscle endurance and strength), arm curl (a measure of upper body muscle endurance and strength), chair sit-and-reach (a measure of lower body flexibility), back-scratch test (a measure of upper body flexibility), 6-min walk (a measure of aerobic endurance), and 2-min step-in-place (an alternative test of aerobic endurance). The norms are both age and gender specific, with normative values reported in 5-year increments (Rikli & Jones, 1999b, 2001).

The performance norms for the UG can be used to directly compare the performance of an older adult with others of the same age and gender. As was the case for the TUG, the UG has also been shown to discriminate between physically independent and dependent older adult women (Jones, Rose, & Newsome, 1999). What remains to be investigated, however, is whether the UG can also accurately predict an older adult’s risk for falling. The two primary goals of the present study, therefore, were to determine whether performance on the UG could discriminate between a group of older adults who were recurrent fallers from a group who had not sustained a fall within the previous year and whether the UG would be as sensitive and specific a predictor of falls as the TUG when applied to a population of community-residing older adults.

**Methods**

A total of 134 community-dwelling older adults (age 60–90 years), 112 women and 22 men, were selected from a larger sample of older adults (N = 209) who completed a fall-risk-reduction screening conducted at their local senior center. Participants completed a comprehensive assessment of their balance and mobility that took approximately 1 hr. In addition to completing both timed up-and-go tests, participants completed the remaining five items of the Senior Fitness Test (Rikli & Jones, 2001), the Berg Balance Scale (Berg et al., 1992), a 50-ft-walk test performed at a preferred and maximum speed, and five additional tests of balance using computerized posturography. The tests performed on the Balance Master® included sit-to-stand, limits of stability, step, quick turn, modified clinical test of sensory
interaction in balance, and tandem walk. All tests were presented in a random order, with the special provision that the two up-and-go tests not be performed consecutively or before any test likely to induce fatigue (i.e., chair stand, 2-min step-in-place) to minimize any test-order effects. Although participants completed a comprehensive fall-risk assessment, only their performance on the TUG (Podsiadlo & Richardson, 1991), and UG (Rikli & Jones, 1999a) are addressed in this article.

In order to be eligible for the study, all participants were required to be 60 years or older, living independently in the community, able to walk at least 50 ft with or without the use of an assistive device, and of normal cognitive function. Fall history was obtained by a self-report measure. Both the frequency and the circumstances surrounding each reported fall were recorded by the participant and subsequently confirmed during a verbal follow-up interview with a member of the testing team. For the purposes of this study, a fall was defined as an event resulting in an individual unintentionally coming to rest on the ground, floor, or other lower level, not as the result of a major intrinsic event or overwhelming hazard (Tinetti, Speechley, & Ginter, 1988).

In addition to the general eligibility criteria established for the study, older adults were excluded from the study if they had fallen on only one occasion during the previous year or reported any medical diagnosis that could have accounted for a fall. Exclusionary diagnoses included cerebrovascular accidents, chronic and progressive neurological disorders (e.g., amyotrophic lateral sclerosis, Parkinson’s disease, multiple sclerosis), transient ischemic attacks, cardiovascular problems, and severe musculoskeletal disorders (e.g., rheumatoid arthritis, osteoarthritis). Medications listed on the health and physical activity questionnaire were also reviewed for the purpose of identifying any medications known to increase fall risk or cause adverse side effects likely to affect balance and mobility (e.g., psychotropics, benzodiazepines, diuretics).

The decision to exclude older adults who had sustained only one fall during the previous year was made in order to maximize the identification of “true” fallers. It has been previously argued that recurrent falls might be more predictable and associated with a physiological predisposition to falling or more serious consequences than a single, isolated fall event (Chandler, 2000; Nevitt, Cummings, Kidd, & Black, 1989). Participants meeting both the general and the specific eligibility criteria for the present study included 71 older adults (mean age 76.6 years, $SD = 6.2$, range 63–89) with no history of falls during the previous year and 63 older adults (mean age 78.1 years, $SD = 6.5$, range 62–90) with a history of two or more falls ($M = 2.8$, $SD = 1.4$) within the same time period. A health and physical activity questionnaire was used to obtain important demographic and health information, as well as each participant’s fall history. Participant characteristics are summarized in Table 1.

Before testing day, participants signed up for a 1-hr test session and were asked to complete a medical health- and activity-history questionnaire. They were also required to obtain a medical release from their primary-care physician in order to participate in the screening. Participants returned their completed forms at the time of their scheduled appointment. They then signed an informed-consent form and completed a balance-self-confidence scale before performing physical assessments associated with the fall-risk-reduction screening. The TUG and UG were performed in a counterbalanced and nonconsecutive order by the same tester, before
### Table 1  Demographics of Older Adults With No History of Falls and Those With a History of Recurrent Falls

<table>
<thead>
<tr>
<th></th>
<th>Nonfallers (n = 71)</th>
<th>Fallers (n = 63)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>76.7</td>
<td>78.4</td>
</tr>
<tr>
<td>$SD$</td>
<td>6.2</td>
<td>6.5</td>
</tr>
<tr>
<td>range</td>
<td>63–89</td>
<td>62–90</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>female</td>
<td>59</td>
<td>53</td>
</tr>
<tr>
<td><strong>Height (in.)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>63.8</td>
<td>64.0</td>
</tr>
<tr>
<td>$SD$</td>
<td>3.1</td>
<td>3.5</td>
</tr>
<tr>
<td>range</td>
<td>56–71</td>
<td>57–76</td>
</tr>
<tr>
<td><strong>Weight (lb)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>147.7</td>
<td>157.1</td>
</tr>
<tr>
<td>$SD$</td>
<td>30.0</td>
<td>36.3</td>
</tr>
<tr>
<td>range</td>
<td>100–260</td>
<td>86–265</td>
</tr>
<tr>
<td><strong>Number of diagnoses</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–1</td>
<td>25</td>
<td>8</td>
</tr>
<tr>
<td>2–3</td>
<td>38</td>
<td>32</td>
</tr>
<tr>
<td>&gt;4</td>
<td>8</td>
<td>23</td>
</tr>
<tr>
<td><strong>Number of Prescriptions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–1</td>
<td>26</td>
<td>12</td>
</tr>
<tr>
<td>2–3</td>
<td>27</td>
<td>25</td>
</tr>
<tr>
<td>&gt;4</td>
<td>18</td>
<td>26</td>
</tr>
<tr>
<td><strong>Assistive Device</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>none</td>
<td>71</td>
<td>32</td>
</tr>
<tr>
<td>cane</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>walker</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>*<em>Up-and-go (s)</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>7.1</td>
<td>10.6</td>
</tr>
<tr>
<td>$SD$</td>
<td>1.4</td>
<td>3.0</td>
</tr>
<tr>
<td>range</td>
<td>4.6–11.7</td>
<td>6.5–11.9</td>
</tr>
<tr>
<td><strong>Timed up-and-go (s)</strong>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>8.2</td>
<td>12.3</td>
</tr>
<tr>
<td>$SD$</td>
<td>1.8</td>
<td>3.9</td>
</tr>
<tr>
<td>range</td>
<td>5.0–14.1</td>
<td>7.0–26.9</td>
</tr>
</tbody>
</table>

* $p < .01$ between groups. ** $p < .001$. 
the performance of test items that might have caused fatigue (i.e., chair stand or 2-
min step-in-place).

The TUG required the participant to rise from a seated position, walk 3 m
(9.84 ft) as quickly and as safely as possible to a line marked on the floor, turn
around, walk back, and sit down in the shortest time possible. Timing began with
a verbal signal to “go” and ended when the participant’s buttocks touched the seat.
Participants were permitted to use their hands to rise from or lower themselves onto
the chair. An assistive device could also be used to complete the test. Test
procedures for the shorter test (UG) were identical, with the exception that the
participant walked quickly around a cone placed at an 8-ft mark on the floor en route
back to the chair instead of crossing a taped line marked at 10 ft on the floor. All
participants received a verbal and visual demonstration of the test and one practice
trial before performing two test trials on both the TUG and the UG. A brief rest
interval was provided between the first and second trials on both tests. Because of the
different scoring protocols associated with each test, the mean time was recorded
for the two trials on the TUG, whereas the fastest time was recorded for the UG.

In order to determine whether there were significant performance differences
between community-dwelling older adults with no history of falls and those who
had sustained multiple falls during the previous year, a univariate analysis was
conducted for both the UG and the TUG. The sensitivity and specificity of the UG
to predict the probability of falls in the same two groups was then examined using
logistic-regression analysis. For the purpose of comparison, sensitivity and speci-
ficity were also calculated for the TUG using logistic-regression analysis. Cutoff
scores for the UG and TUG that maximized each test’s sensitivity (i.e., percentage
of time the test correctly classified older adults as fallers) and specificity (i.e.,
percentage of time the test correctly classified older adults as nonfallers) were
calculated using the same procedure. All statistical analyses were performed using
SPSS, version 10.0 (SPSS, 1999).

Results

In order to determine whether the two groups’ performance on the UG and TUG
were significantly different, two one-way ANOVAs were conducted. The results in
Table 1 indicate that there were significant performance differences between the
groups for both tests, $F(1, 132) = 75.01, p < .01$, for the UG and $F(1, 132) = 63.20,$
$p < .001$ for the TUG. For the UG, the mean performance score for the nonfaller
group was 7.1 s ($SD = 1.4$) and for the faller group was 10.6 s ($SD = 3.0$). On the TUG,
the mean performance scores were 8.2 s ($SD = 1.8$) and 12.3 s ($SD = 3.9$) for the
nonfaller and faller groups, respectively.

The sensitivity and specificity of the TUG and the UG were also examined
using logistic-regression analyses. Multiple cutoff levels that had a predicted
probability of .50 or higher and maximized both the sensitivity and specificity of the
test were examined for the purpose of identifying a performance-time score that
most accurately classified recurrent fallers and nonfallers. As shown in Table 2,
older adults who required 8.5 s or longer to complete the UG were classified as
fallers, with an overall prediction rate of 82%. The specificity of the test was good,
with a total of 61 out of 71 nonfallers, or 86%, being correctly identified, whereas
the sensitivity of the test was slightly lower, with 49 of the 63 participants, or 78%, being correctly classified as fallers. By way of comparison, the sensitivity and specificity of the TUG test were also calculated for the sample of older adults in the present study. A cutoff score of 10 s on the TUG yielded an overall prediction rate of 80% in the present study. The sensitivity of the test using this cutoff value was 71%, and the specificity was 89%. Had the cutoff score of 13.5 s established by Shumway-Cook and colleagues (2000) been used, however, the overall prediction rate using the present participant sample would have been considerably lower, at 64%. This is in stark contrast to the overall prediction rate of 90% reported in the earlier study. In the present study, only 19 of the 63 actual fallers required 13.5 s or longer to complete the TUG, yielding a sensitivity of 30%. Conversely, the specificity of the test remained high, with 70 of the 71 nonfallers, or 99%, being correctly identified (refer to Table 2).

### Discussion

The two primary research questions addressed in this study were (a) whether there were significant performance differences on the UG test between a group of community-residing older adults with a history of falls and a group with no such history and (b) whether the UG test would be as sensitive and specific a predictor of falls as the TUG when used to screen the same group of older adults. As previously demonstrated with the TUG (Gunter et al., 2000; Shumway-Cook et al., 2000), significant differences in performance were evident between a group of recurrent fallers and nonfallers tested on the shorter version UG. Using the national norms established by Rikli and Jones (1999b) to compare the performance of the two groups studied, it is interesting to note that although the performance of the nonfaller group was within the “normal” range of scores established for men and women between 75 and 79 years of age, the performance of the faller group was well below average. The mean UG score of 10.6 s recorded for this group placed them below the 5th percentile for men and women in not only the same-age group but also in the 80- to 84-year age group. The mean performance times recorded by this group would only have placed them at the low end of the normal performance range (for

<table>
<thead>
<tr>
<th>Test</th>
<th>Cutoff value</th>
<th>Specificity, nonfallers ($n = 71$)</th>
<th>Sensitivity, fallers ($n = 63$)</th>
<th>Overall prediction</th>
<th>Predicted probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>UG</td>
<td>8.5</td>
<td>61/71 (86%)</td>
<td>49/63 (78%)</td>
<td>82%</td>
<td>.60</td>
</tr>
<tr>
<td>TUG</td>
<td>10.0</td>
<td>63/71 (89%)</td>
<td>45/63 (71%)</td>
<td>80%</td>
<td>.60</td>
</tr>
</tbody>
</table>
women only) had they been between 90 and 94 years old. These results clearly indicate the degree to which the functional mobility of this faller group was compromised.

With respect to the second research question addressed in the present study, the results indicated that the UG and the TUG were both able to predict which older adults were most likely to be fallers, using a cutoff score of 8.5 s in the case of the UG and a 10.0-s cutoff value for the TUG. Although the sensitivity of the UG was a little higher than that of the TUG (76% vs. 71%), the UG demonstrated slightly lower specificity, with 86% of the nonfallers being correctly identified and 89% of fallers being correctly identified using the TUG. In light of the fact that maximizing sensitivity at the cost of minimizing specificity is desirable when the consequence of a false negative (i.e., falsely identifying a faller as a nonfaller) exceeds the consequence of a false positive (i.e., falsely identifying a nonfaller as a faller), the higher sensitivity of the UG is considered the more important difference between the two tests of functional mobility. These results would suggest that the UG is a comparable predictor of fall risk among community-residing older adults when compared with the TUG.

The results of the present study are in stark contrast to those obtained by Shumway-Cook and colleagues (2000), who reported high sensitivity (80%) and specificity (100%) values for the TUG using a cutoff score of 13.5 s. As indicated earlier in the Results section, had we adopted the same cutoff score in the present study we would have erroneously classified 45 of the 63 participants known to be recurrent fallers as nonfallers. The large differences in cutoff values can be attributed to the large differences in sample size (139 versus 30) and the characteristics of the participants who participated in the two studies. Although the mean ages of the faller and nonfaller groups in the present study were not significantly different (i.e., 76.7 years vs. 78.4 years), that was not the case for the earlier study, in which the age of the faller group was significantly higher than that of the nonfaller group. Moreover, the participants in the Shumway-Cook et al. study were considerably more frail, as indicated by the considerably longer mean time required to complete the test and the fact that 12 of the 15 fallers in that study required some form of assistive device to perform the test. Only 32 of the 63 fallers in the present study reported using an assistive device on a daily basis, and none used a device for the purpose of completing either test. Finally, the inclusion criteria were also considerably more stringent in the earlier study—only older adults who had fallen at least twice within the past 6 months were included in the faller group. It is perhaps no surprise, therefore, that the overall prediction rate reported in that study was so high (i.e., 90%).

Although it might be argued that the findings of the present study are less meaningful because we chose not to include a group of older adults who had sustained only one as opposed to multiple falls in the previous year, our decision was guided by previous arguments in the literature that recurrent falls might be not only more predictive but also more likely to signal a true physiological disposition to falling (Chandler, 2000; Nevitt et al., 1989). In reviewing our own database comprising more than 500 community-residing older adults who have completed comprehensive balance and mobility screenings within the past 3 years, of the 109 adults who reported sustaining a single fall during the previous year, approximately 50% of the older adults tested were considered to be at low risk for falls based on
their performance on multiple clinical tests of balance and mobility, as well as impairment-based tests using computerized posturography. Only 12 older adults who had reported a single fall during the previous year were categorized at high risk after completing the comprehensive assessment. These additional data would appear to support previous arguments that sustaining a single fall in a 12-month period is not a reliable indicator of actual fall risk.

Although it is important that clinicians be able to identify individuals who are in need of immediate intervention, using a cutoff score assumed to predict the probability of falls from a single test of balance might not constitute the best decision-making method. One alternative and perhaps more appropriate method that could be adopted by developers of screening tests would be to calculate threshold values or criterion reference points that identify when an individual is at risk for loss of functional mobility or independence. Clinicians could then use these values to guide their recommendations for further treatment. Such criterion reference points have been established for each of the physical tests that make up the Senior Fitness Test (Rikli & Jones, 2001). For the UG, for example, criterion reference points of 8.9 and 8.8 s were established for men and women, respectively. These scores mark the upper boundary of the zone for those at risk for loss of functional mobility or independence. It is interesting to note, however, how very close these values are to the fall-risk cutoff value of 8.5 s established in the present study. Given that a clinician’s goal should be to identify older adults in need of immediate intervention before they begin falling, the use of criterion performance scores might constitute a better method for making appropriate treatment recommendations.

Although it is well known that the etiology of falls is multifactorial in nature, simple tests such as the UG represent a valid and effective method of discriminating between adults of different ages, activity levels, and levels of functional independence. The results of this study further demonstrate the ability of the UG not only to discriminate between older adults who are experiencing multiple falls and those who are not but also to predict which older adults are at high risk for falls with a good degree of specificity and sensitivity. In light of the fact that age- and gender-matched national norms and criterion reference points have been established for the UG for older adults ranging in age from 60 and 94 years, as well as the fact that the test is more easily and safely administered in smaller spaces, we would recommend the adoption of this functional mobility test by practitioners as a valid and effective method of screening for disability and fall risk in the older adult population.

Acknowledgment
This work was funded, in part, by Archstone Foundation Grant No. 98-02-14.

References


