Reliability, Validity and Minimal Detectable Change in the Timed Up and Go and Five Times Sit to Stand Tests in Older Adults with Early Cognitive Loss

Jennifer Blackwood*

Abstract

Background: Five times sit to stand test (FTSTS) and Timed Up and Go (TUG) have been used in clinical settings as a measure of physical performance in older adults with and without cognitive dysfunction. The validity and reliability of these measures have been established in older adults with differing diagnoses, but not in those with early cognitive loss. The purpose of this study was to evaluate the validity, reliability and minimal detectable change of FTSTS and TUG in older adults with early cognitive loss.

Methods: Performance on the FTSTS and TUG was assessed in 26 older adults. Test-retest reliability was examined using ICC and absolute (SEM) reliability as well as the MDC95. Pearson’s correlation coefficient was used to examine relationships between the measures and gait speed, to determine construct validity. Bland-Altman plots were constructed to assess systematic bias.

Results: FTSTS had high test-retest reliability (ICC = 0.89), small SEM (1.20 s), and the MDC95 was 3.54 s. TUG had high test-retest reliability (ICC = 0.81), small SEM (1.60 s), and the MDC95 was 5.37 s. Correlation coefficients between the measures and gait speed indicates that FTSTS and TUG are valid measures of dynamic balance in older adults with early cognitive loss.

Conclusions: To be considered real change beyond measurement error, change in FTSTS performance should exceed 3.54 seconds, TUG should exceed 5.37 seconds.

Keywords

Community dwelling older adults; Early cognitive loss; Chair rise test; Balance

Introduction

Older adults (age 65+) with impaired cognitive function are twice as likely to fall as cognitively intact peers and that risk increases if they have a history of falling or impaired mobility [1]. Intact cognitive function is necessary to maintain stability and balance during mobility and functional activities [2]. Deficits in cognitive function have been reported to be an independent risk factor for injury or multiple falls [3]. The influence of cognitive performance on mobility is present throughout the spectrum of cognitive decline, including early in the decline, as impairments in balance and gait have been reported in those with early cognitive loss [4]. Associations between cognitive function and balance have been reported in older adults with diagnosed mild cognitive impairment, but have also been found in those without diagnosed cognitive impairment further highlighting the underlying role that cognitive function has on safe mobility in older adults [5,6].

Several measures of functional mobility exist which are used by clinicians to determine risk of falling and mobility impairments. One tool which is easily performed, the Five Times Sit to Stand test (FTSTS) assesses balance during transitional movements and can be used across many different settings [7]. FTSTS has established reliability and validity in older adults with COPD Parkinson’s disease, cardiovascular disease and in community dwelling older women [8-11]. Longer FTSTS completion times are associated with increased disability and falls and have been reported in those with moderate to severe global cognitive dysfunction [7,12]. FTSTS has been suggested for use in predicting falls in older adults and for predicting cognitive decline in older women; however the validity and reliability of this measure has not been reported in those with early cognitive impairment [7,12]. Similarly, the Timed Up and Go (TUG) is a quick objective measures to assess dynamic balance and mobility. TUG has established reliability in those with Alzheimer’s disease and both reliability and validity in community dwelling older adults [13-15]. Longer TUG completion times have been associated with poorer performance on measures of global cognition, verbal fluency, working memory, processing speed, and executive function in older adults with mild cognitive impairment [5].

Detection of mobility impairment in older adults with early changes in cognition offers the opportunity for timely treatment and initiation of preventative measures to decrease risk of injury, yet tools commonly used to screen functional mobility and falls risk, like the FTSTS and TUG do not have established reliability and validity in this population. Therefore, the aim of this study was to evaluate the validity, reliability and minimal detectable change of the FTSTS and TUG in community dwelling older adults with early cognitive loss. It was hypothesized that that the psychometric properties of the FTSTS and TUG are comparable to those found in healthy older adults without cognitive loss.

Methods

Non-institutionalized older adults (age 65+) were recruited through presentations and written advertisement at three community senior centers. Participants were included in the study if they met criteria and screened positive for having early cognitive loss: self-reported difficulty with cognitive function, a Montreal Cognitive Assessment (MoCA) score of <26, impaired executive function as indicated by a prolonged Trail Making Test part B (TMT-B) score (>77 seconds) and deficits in visual spatial ability as indicated by impaired Cube Copying Test performance. Those who met the inclusion and exclusion criteria went through initial testing. Other inclusion criteria were: good vision with or without prescriptive lenses, independent with mobility (use of assistive devices was acceptable), English speaking, and able to independently provide consent for
Participants were excluded if they had surgery or had been hospitalized six months prior to testing, were currently receiving mobility related rehabilitation services, were unable to meet the time commitment for the study, or if they had a history of being diagnosed with moderate to severe cognitive dysfunction (dementia, Alzheimer’s disease, Parkinson’s disease), cerebrovascular accident, head trauma, or other traumatic brain injury. Human subject institutional review board approval was obtained.

To assure consistency between examiners, a one-day training session covering the testing protocol and administration of all measures was completed. Examiners were trained on the use of the cognitive measures using the administration and scoring resources available at the time of testing [16]. During the training session examiners practiced administering and interpreting all assessments on four volunteers (age range 58-65).

After providing informed consent, participants completed a medical interview for comorbidities, medication usage, and the number of falls in the six months prior to testing [17]. History of falling was by self-report and a fall was defined as the unintentional loss of balance that led to an unexpected change of position. For further clarification, examples of a fall were provided including falls which did not result in hitting the ground, but other areas (i.e. falling into a chair or a wall). Other information gathered included age and education level. All testing occurred in a private room at the senior center and standardized instructions were used for all measures. Cognitive assessments used to confirm the presence of early cognitive dysfunction consisted of the TMT-B, the Cube Copying Test, and the MoCA and were performed in that order. The physical measures were completed afterwards. Each physical measure was first explained to the participants followed by a demonstration of the task. Physical measures were completed twice with the average of trials recorded. One examiner performed the testing session at each trial (initial and 6 week follow up).

**Trail making test part B**

TMT-B assessed executive function [18-20]. In this timed paper and pencil test, participants draw a continuous line creating the trail without lifting the pencil connecting randomly placed letters (A-M) and numbers (1-12) in ascending alternating order from number to letter (e.g., 1 to A to 2 to B to 3 etc.) until all numbers and letters are used. Demonstration of how to complete the TMT-B was provided and participants practiced on a different shortened version of the TMT-B. Timing began after the test paper was turned over and ended after correctly connecting all letters and numbers on the paper. The maximum time allotted for the TMT-B is 300 seconds at which the test is stopped and that time is recorded [20]. Time was recorded as a continuous score. TMT-B performance has been reported to be associated with age and education and a score of >77 seconds was considered to be impaired performance [21].

**Cube-copying test**

The Cube-Copying Test is a measure of visual spatial ability which requires the participants to copy the Necker Cube. In this test, participants were instructed to make a copy of the cube below the drawing, with no time limitation for completion. A correctly executed ‘non-impaired’ Cube-Copying Test was given a score of ‘1’ if the drawing was three-dimensional, had the correct number, placement, and intersection of lines, and had all lines parallel to each other and similar in length. If errors on any of those criteria were found, then a score of ‘0’ was assigned indicating an ‘impaired’ drawing. Scoring for the Cube-Copying Tests was performed initially by one examiner and then confirmed by group consensus by two other external researchers with training in interpretation of these measures. Poor performance on the Cube-Copying Test has been associated with impaired cognitive function (TMT-B), impaired motor ability and falls [22,23].

**Montreal cognitive assessment**

As a global measure of early cognitive loss, MoCA assesses several cognitive domains including attention and concentration, executive function, memory, language, visuocunstructional skills, conceptual thinking, calculations, and orientation [16]. Scoring of each of the items constitute the total MoCA score (ranging from 0-30) with lower scores indicating impairment in cognitive function. A cutoff score of 26 has a sensitivity of 90% and a specificity of 78% for detecting mild cognitive impairment and was, therefore, used as the cut point for detection of early cognitive loss.

**Five times sit to stand (FTSTS)**

FTSTS consists of recording the amount of time it takes for a person to go from sitting in a supportive chair (47.5 cm in height) to a full standing position five successive times as quickly as possible without the use of arm [7]. Participants were given a 1-minute rest break between trials. Timing began when the participant was told to ‘go’ and when the participant’s buttock contacted the chair the fifth time.

**Timed up and go**

For this test, participants stood up from a chair, walked 3 meters at a quick pace to a line clearly marked on the floor, turned around after crossing over the line, walked back to the chair, and sat down [24]. Timing of this task began from the moment the subject was told to ‘go’ and ended with sitting back down in the chair.

**Gait speed**

Usual gait speed was recorded as the amount of time required to walk 3.048 meters (10 feet) at the participant’s normal walking speed. To allow for acceleration and deceleration, 1.524 meters (5 feet) of space was provided on each side of the marked distance. Speeds were later converted to meters/second in the analyses.

**Statistical Analysis**

**Sample size**

The sample size was estimated on the basis of a power of 0.8 and an alpha level of 0.05 (2-tailed). For the test-retest reliability analysis, the sample size was estimated based on: 2 trails of testing by the same rater with a minimal acceptable ICC value of 0.70 [25]. At least 19 participants would be required for the test-retest reliability. For the construct validity, a large effect size (ρ=0.50) was chosen to estimate the sample size using G*Power 3.1 [26]. At least 26 participants would be required for the analysis of construct validity. Descriptive statistics including means, standard deviations, ranges, and percentages were calculated. The distributions of data were assessed using the Shapiro-Wilk tests. Statistical significance was set at p< 0.05 and analysis was performed using IBM-SPSS 24 (IBM Corp., Armonk, NY).

**Reliability**

Test-retest reliability was analyzed using the intraclass correlation
coefficient (ICC_{2,1}) with absolute agreement, single measure and a 95% CI [27]. Good reliability is indicated by an ICC value greater than 0.75, however an ICC value greater than 0.90 is recommended [27]. The standard error of the measurement (SEM) was used to assess consistency of scores from each trial for each participant. The SEM indicates the stability of the score when a measure is repeated. Smaller SEM numbers indicate a more reliable measurement. The following equation was used to calculate the SEM: SD x \sqrt{(1-ICC)}, where the SD is the pooled standard deviation of the first and second FTSTS and TUG measurements and the ICC is the correlation coefficient from the test-retest reliability analysis [27].

**Minimal detectable change**

The minimum amount of change in a score which is believed to not be the result of measurement error nor within-subject variability is known as the Minimal Detectable Change (MDC) [27]. The MDC at 95% a confidence interval (MDC95) was used to interpret changes in FTSTS and TUG scores. MDC95 was calculated using the following equation: MDC95=SEM x \sqrt{2}. A change in scores which is greater than the MDC95 indicates a true change in a person’s performance with 95% confidence [27].

**Validity**

Pearson’s correlation was used to examine construct validity of the scores on each of the two measures (FTSTS and TUG) with gait speed. Correlation strength was determined using those reported by Portney and Watkins: little-none (r<0.25), poor (r=0.25-0.50), moderate (r=0.50-0.75) and good-excellent (r>0.75) [27]. Bland-Altman plots of each participant’s mean FTSTS or TUG score plotted against their FTSTS or TUG difference score (trial 1- trial 2) were constructed to display the spread of difference scores about the mean difference score. The Bland-Altman plots also displayed the 95% constructed to display the spread of difference score. The Bland-Altman plots of each participant’s mean FTSTS or TUG score plotted against their FTSTS or TUG difference score (trial 1- trial 2) were constructed to display the spread of difference scores about the mean difference score. The Bland-Altman plots also displayed the 95% confidence interval (95% LOA) which represents the expected range of difference scores across trials of the measure (FTSTS or TUG) [27]. The 95% LOA was calculated as the difference in mean scores of the measure (FTSTS or TUG) across trials + SD x 1.96, with the SD as the standard deviation of the difference scores.

**Results**

A convenience sample of 32 older adults self-selected to participate in the study. In the screening process, two participants were excluded because their TMT-B scores (45.01, 57.43 s) and Cube-Copying Test performance were not impaired despite poorer performance on the MoCA (scores of 23, 24) and self-report of cognitive dysfunction [21-23]. Four additional participants were excluded from the analyses; one due to having a history of a stroke, one with a diagnosis of Parkinson’s disease, and two did not return for the second testing session. Twenty-six older adults (mean age in years: 76.23 (6.06), 65.4% female) with early cognitive loss [mean MoCA= 22.73 (2.37), TMT-B= 102.90 (37.40)] met inclusion criteria, provided informed consent, and completed testing. Characteristics of the sample can be found in Table 1. Mean scores, in seconds, of trial 1 of the FTSTS and TUG were 14.55 (3.76) and 10.89 (4.48) respectively. FTSTS and TUG scores and ranges by trial, SEM, ICC_{2,1}, 95% CI, and MDC_{95} are found in Table 2. Moderate construct validity was found for both measures. FTSTS performance was inversely associated (r=-0.53, p<0.05) with gait speed indicating that longer FTSTS completion time was associated with slower walking speed in those with early cognitive loss. Performance on the TUG was inversely associated with gait speed (r=-0.58, p<0.05) indicating that as gait speed decreased, TUG scores completion times increased.

**Reliability and minimal detectable change**

Mean FTSTS times per trial were similar as the ICC_{2,1} of 0.89 indicated high test-retest reliability. The SEM for FTSTS was small and calculated as 1.2 seconds. The MDC95 was 3.54 seconds. For the TUG, the ICC_{2,1} was 0.81 indicating good test-retest reliability and a small SEM of 1.6 seconds was calculated. The MDC95 of TUG was 5.37 seconds. Bland-Altman plots of the FTSTS and TUG scores

Table 1: Characteristics of older adults with early cognitive loss (n=26).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean (SD) or %</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic/ Anthropometric</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>76.23 (6.06)</td>
<td>66-91</td>
</tr>
<tr>
<td>Gender % Female</td>
<td>65.4</td>
<td></td>
</tr>
<tr>
<td>Education, % with a Bachelor’s degree or higher</td>
<td>42.3</td>
<td></td>
</tr>
<tr>
<td>Prescription Medications taken daily</td>
<td>7.28 (5.32)</td>
<td>1-24</td>
</tr>
<tr>
<td>Body mass index (Kg/m²)</td>
<td>27.15 (4.34)</td>
<td>15.6-38.4</td>
</tr>
<tr>
<td>Comorbidity Index</td>
<td>3.53 (1.79)</td>
<td>1-8</td>
</tr>
<tr>
<td>Physical Mobility &amp; Cognitive Function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greater than 1 fall in the past 6 months, %</td>
<td>46.2</td>
<td></td>
</tr>
<tr>
<td>Gait Speed (m/s)</td>
<td>0.96 (0.27)</td>
<td>0.45-1.45</td>
</tr>
<tr>
<td>MoCA</td>
<td>22.73 (2.37)</td>
<td>17-25</td>
</tr>
<tr>
<td>TMT-B</td>
<td>102.90 (37.40)</td>
<td>50.0-175.5</td>
</tr>
</tbody>
</table>

MoCA: Montreal Cognitive Assessment test; TMT-B: Trail Making Test part B

Table 2: Test-retest reliability, SEM, and MDC95 of the five times sit to stand and timed up and go tests in older adults with early cognitive loss (n=26).

<table>
<thead>
<tr>
<th></th>
<th>Mean trial 1 (SD)</th>
<th>Mean trial 2 (SD)</th>
<th>ICC_{2,1}</th>
<th>95% CI</th>
<th>SEM</th>
<th>MDC_{95}</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTSTS</td>
<td>14.55 (3.76)</td>
<td>14.10 (3.18)</td>
<td>0.89</td>
<td>0.72, 0.95</td>
<td>1.20</td>
<td>3.54</td>
</tr>
<tr>
<td>TUG</td>
<td>10.89 (4.47)</td>
<td>10.51 (2.94)</td>
<td>0.81</td>
<td>0.57, 0.92</td>
<td>1.60</td>
<td>5.37</td>
</tr>
</tbody>
</table>

FTSTS: Five Times Sit to Stand, TUG: Timed Up and Go

Citation: Blackwood J (2017) Reliability, Validity and Minimal Detectable Change in the Timed Up and Go and Five Times Sit to Stand Tests in Older Adults with Early Cognitive Loss. J Physiother Rehabil 1:1.
commonly used in older adults with early cognitive dysfunction. This study of functional mobility and balance are reported in a group of community dwelling older adults including those with COPD Parkinson’s disease, cardiovascular disease and in community dwelling older women but has not been reported in older adults with cognitive impairment [8-11]. In older females without cognitive impairment, test-retest reliability was better (ICC$_{2,1} = 0.95$) than what was found in this study (ICC$_{2,1} = 0.89$), however, the reliability is still acceptable and is better than those with Parkinson’s disease (ICC$_{2,1} = 0.76$) and those with cardiovascular disease (ICC$_{2,1} = 0.87$) [8,9,11]. In comparison, the MDC$_{95}$ in those without cognitive impairment was less (2.5 – 3.12 s) as compared to the minimal amount of change required in those with cognitive dysfunction reported here (3.54 s) indicating that for those with early cognitive changes, FTSTS times must exceed those reported in cognitively intact older adults to achieve the minimum difference which exceeds measurement error.

Older adults, a linear relationship exists between falls risk or falls incidence and the presence and extent of cognitive dysfunction, with a greater risk and incidence of falls with more impaired cognition. As a measure of mobility and falls risk, performance on the FTSTS has been reported to be associated with lower levels of global cognitive function as measured with the Mini Mental State Exam; and the Short Portable Mental Status Questionnaire measures commonly used to assess for moderate to severe dementia [12,28-30]. However, FTSTS performance of the participants of this study was, on average, better than those with more impaired cognitive function (14.55 s vs. 17.79 s) [12]. As FTSTS performance varies among those with differing levels of cognitive function, the minimum amount of change in a score which is believed to not be the result of measurement error nor will within-subject variability likely differ throughout all levels of cognitive impairment. Therefore, the MDC reported in this paper should be considered for older adults with early cognitive loss as FTSTS scores and the MDC$_{95}$ would likely differ for those with a greater degree of cognitive dysfunction.

TUG has previously been reported to have excellent test-retest reliability in community dwelling elderly people with ICC values ranging from 0.97 to 0.99, however reliability was reported to be much lower in those with cognitive impairment [ICC = 0.56 [13-14,24]. In this study of adults with early cognitive loss, the test-retest reliability was much better (ICC = 0.81) with a small SEM of 1.6 seconds indicating that TUG is a reliable tool to use in this population, even with a window of 6 weeks between testing. Additionally, the MDC$_{95}$ reported by Ries et al. in 2009 in older adults with Alzheimer’s Disease was smaller (4.09 seconds) than what was found in this study (5.37 seconds) indicating that in those with early cognitive loss, the amount of improvement on the TUG must be greater than what would be expected for those with more advanced cognitive loss. Performance on the MoCA, TMT-B, and Cube-Copying Test along with self-reported compliant of difficulties with cognitive function comprised the criteria for identification of early cognitive loss. As a screening tool, MoCA has good sensitivity and specificity for identification of early cognitive loss, however, the use of one measure to indicate early cognitive loss limits the ability to detect cognitive impairments reported to be associated with physical mobility. Therefore, measures which examine cognitive constructs associated with falls and balance impairments were included to identify the presence of early cognitive loss which adds to the strength of the classification of the sample.

Our study was limited in that the population was a convenience sample of educated white older adults from one geographic location who self-selected to participate which does not permit generalizability of study results to those of different education levels or race. Reporting of falls history via a six-month recall is a limitation of the
study in that falls may have been under-reported compared to using a falls diary. We sought to include participants who did not have significant mobility limitations and therefore, performance on the physical performance measures reflected the characteristics of this group which may not be consistent with community dwelling older adults with more progressive physical or cognitive impairments. Future studies of FTSTS measurement error and minimum change should consider recruiting adults with more advanced cognitive loss, those with a greater number of comorbidities, and different settings (e.g. skilled nursing). Additional psychometric properties such as the minimal clinically important difference of the FTSTS and TUG should be considered in future studies of older adults with more advanced cognitive impairment.

Conclusion

TUG and FTSTS are reliable and valid measures of functional mobility for community dwelling older adults with early cognitive loss. The MDC can be used by clinicians to examine the outcomes of balance and mobility related treatments to address balance impairments in older adults with cognitive impairment. Due to the progressive nature of cognitive decline commonly experienced by older adults, appropriate identification and treatment of physical limitations found in early cognitive loss may contribute to a higher level of functional mobility, independence, and effective management of comorbidities.

Acknowledgment

The author would like to thank the older adults who volunteered to participate in this study and Heather Hess and Megan Fuson for assistance with data collection.

Declarations of Interest

The author reports no declarations of interest.

References


Author Affiliations

Department of Physical Therapy, School of Health Professions and Studies, University of Michigan-Flint, Michigan, United State of America