International Journal of Gerontology 10 (2016) 37–42

Contents lists available at ScienceDirect

International Journal of Gerontology

journal homepage: www.ijge-online.com

Original Article

# Influence of Procedural Factors on the Reliability and Performance of the Timed Up-and-go Test in Older Adults☆

Szu-Ping Lee [a *], Janet Dufek [b], Robbin Hickman [a], Sue Schuerman [a]

[a] Department of Physical Therapy, University of Nevada, Las Vegas, NV, USA, [b] Department of Kinesiology and Nutritional Science, University of Nevada, Las Vegas, NV, USA

ARTICLE INFO

*Article history:*
Received 8 July 2015
Received in revised form
2 September 2015
Accepted 29 October 2015
Available online 8 March 2016

*Keywords:*
assessment,
mobility,
outcome measure,
stature

SUMMARY

*Background:* "Timed up-and-go" (TUG) test is a clinical assessment commonly used to quantify functional mobility and fall risk in older adults. While the concept of the test is simple, procedural variations can significantly impact the reliability and the performance. The purpose of the study was to determine the influence of specific procedural factors on the test–retest reliability and TUG test performance in older adults.

*Methods:* Adults over the age of 60 years ($N$ = 83, mean age, 69.3 ± 6.9 years; range, 60–91 years) participated. The procedural factors examined included: (1) timing method (handheld stopwatch vs. load-based timing); (2) distances of test (3 m, 6 m, and 9 m); and (3) seat height (standard vs. individual specific). Test–retest reliability obtained from each combination of timing and distance was evaluated. The interaction of seat-height settings and the stature of the participant on TUG performance was investigated by comparing the TUG performance of short and tall participants in the two seat-height settings.

*Results:* Timing method and walking distance modestly influenced the TUG test reliability. The current standard procedure (stopwatch timing and 3-m distance) yielded the lowest but acceptable reliability (intraclass correlation coefficient = 0.887). Taller individuals exhibited significantly better TUG performance when individualized seat heights were used in comparison with the standard lower seat height. The influence of seat height was not as pronounced in shorter individuals.

*Conclusion:* Seat height is an important procedural factor affecting the performance of the TUG test, especially in older adults who are taller. Load-based timing may be used to improve the consistency of the TUG performance assessment.



## 1. Introduction

The "timed up-and-go" (TUG) test is one of the most widely used clinical assessments for mobility performance. This quantitative test[1] has been widely used as the gold standard for assessing functional mobility and fall risks in older adults[2–5]. Over the past 3 decades, the TUG test has been validated and applied in many different populations, including those affected by orthopedic[6], neurological conditions[7–10], cognitive impairments[11], and chronic diseases[12].

While the test is widely used for the geriatric population, some limitations of the test have been identified. For example, in a study testing 1115 older adults, Rockwood et al[11] reported that the TUG test exhibited poor test–retest reliability. In direct contrast to many other studies demonstrating good reliability of the test, this finding showed that it can be a challenge to achieve consistent test results when the test is applied to individuals with a wide range of physical capacities in different testing environments. The authors suggested that one of the causes underlying the observed inconsistency may be the procedural variability in administering the test.

The appeal of the TUG test as a clinical assessment stems from its brevity and the minimal requirements of time and equipment. The test is comprised of timing how long it takes a patient to rise from a seat, walk 3 m, turn around, walk back to the seat, turn around, and return to the seated position. While the procedure is simple, several possible sources for error exist. Firstly, there is no general agreement or specific instruction on how the timing of the

☆ Conflicts of interest: All contributing authors declare that they have no conflicts of interest.

* Correspondence to: Dr Szu-Ping Lee, Department of Physical Therapy, University of Nevada, Las Vegas, 4505 South Maryland Parkway, Box 453029, Las Vegas, NV 89118, USA.
*E-mail addresses:* szu-ping.lee@unlv.edu, szupinglee@gmail.com (S.-P. Lee).

http://dx.doi.org/10.1016/j.ijge.2015.10.003

task should be performed. For example, in Podsiadlo and Richardson's[1] original paper, they described: "… on the word 'go', he is to get up and walk … return to the chair, and sit down again. Either a wrist-watch with a second hand or a stop-watch can be used to time the performance." This instruction is not clear whether the timing should be started at "go", or when the patient initiates the "get up" movement, or when the patient is up fully. Lacking standardized instruction on how to time the performance, clinical testers often have to operationally decide when to start and stop timing. This can be a significant source of error, since older adults usually take much longer and sometimes several tries to complete the "get up" and "sit down" movements. For a walking test that typically takes only 7.1–12.7 seconds to complete[13], small variations in timing can have a large effect on the test reliability. In addition, because the standard walking distance (3 m) and the time to complete the test are short, the effect of slight timing error is augmented. For example, a timing error of 1 second would introduce proportionately greater error to a test that takes 10 seconds to complete (10% error) versus a test that takes 20 seconds to complete (5% error). It is reasonable to believe that a slightly longer walking distance may improve the consistency and reliability of the test.

Lastly, the original testing procedure indicated that "a standard arm chair (approximate seat height of 46 cm)" should be used[1]. However, in clinical practice the test is often administered with any chair available with little concern for how the seat height may influence the TUG performance. This can be problematic as taller individuals may find the chair of choice to be too low and difficult to rise up from, while shorter individuals may find it easier to get up from the same chair. Biomechanically, it has been shown that the required knee extensor moment is significantly greater when a person rises from a lower seated position versus a taller position[14]. As rising up is part of the TUG test, it is important for clinicians to understand the impact of seat height and patient stature on TUG performance. To the best of our knowledge, the influence of seat height on TUG performance has not been evaluated.

The purpose of the study was to evaluate the impact of procedural factors (timing method and walking distance) on the reliability of the TUG test in older adults. Additionally, we investigated the impact of seat height and patient stature on the TUG performance. We hypothesized that an instrumented timing method and longer walking distances would yield greater measurement reliability. We also hypothesized that seat-height settings would influence the TUG performance in older adults.

## 2. Materials and Methods

### 2.1. Patients

A sample of convenience of persons over the age of 60 years ($N = 83$, $M_{age} = 69.3$ years, range, 60–91 years), including 26 men and 57 women, were recruited from local community centers. Participants were enrolled if they were able to walk at least 50 m without help from others. Participants were excluded if they presented with any of the following: (1) inability to understand and follow the verbal instructions given by the investigators regarding the test procedures; (2) injuries which cause pain and/or inability to walk for more than 50 m; or (3) concurrent health conditions that impaired their ability to safely perform light physical activities.

Prior to participation, the objectives, procedures, and risks of the study were explained in detail. Informed consent as approved by the University of Nevada, Las Vegas, Biomedical Institutional Review Board was obtained from each participant.

### 2.2. Instrumentation

An instrumented stool was custom fabricated (Figure 1). The stool composed of a seating surface, a strut system, and a base. The strut system allowed adjustment of the total seat height from 27 cm to 60 cm. A force platform (PS-2142, PASCO Scientific, Roseville, CA, USA) was attached to the base to monitor the magnitude of normal force load. During data collection, force data were streamed digitally via a universal serial bus interface (PS-2100A, PASCO Scientific) to a computer with the acquisition software (DataStudio version 1.9.8r10, PASCO Scientific) at a sampling rate of 100 Hz. The force platform was factory calibrated and zeroed prior to each data collection session. The conventional timing was done by a trained, reliability-proven tester using a standard handheld digital stopwatch sensitive to 0.01 seconds.

### 2.3. Procedure

Data collection was conducted on the campus of the University, or at a local community center. Each participant's sex, age, weight, height, and general health information was obtained. For the purpose of examining the influence of procedural factors on the test–retest reliability, a subset of 15 participants was tested in two sessions on 2 different days. The sessions were between 1 day and 10 days apart. The selection of these participants was based on their willingness and availability to be retested.

Three procedural factors were investigated: (1) timing method (handheld stopwatch vs. load-based timing); (2) walking distance (3 m, 6 m, or 9 m); (3) and seat height (standardized vs. individual specific). The 6-m and 9-m distances were chosen in addition to the standard 3 m because the increases in distance would not require a drastic increase in testing space. Each participant was tested for walking distances (3 conditions) and seat heights (2 conditions) for a total of six combinations. Load-based timing and stopwatch timing were done simultaneously, while three trials were collected from each combination of the distance and seat-height conditions. Preliminary testing was conducted to ensure that our participants, adults older than 60 years, could tolerate the data-collection procedure without developing excessive fatigue. We found that the duration of a complete data collection session would take less than 25 minutes including ample rest time between trials.

**Figure 1.** The instrumented, height-adjustable stool used in the study.