Concurrent validity and minimum detectable change of Senior Fitness Test components: instrumented vs. manual assessment

Megan D. Salvatore1, Alexandra M. Colacino2†, Meghan E. Hess2†, Samuel W. Todd2† and Nathan W. Saunders3*

†These authors contributed equally to this work.
1Department of Physical Therapy faculty, University of Mount Union, USA.
2Department of Human Performance and Sport Business student, University of Mount Union, USA.
3Department of Human Performance and Sport Business faculty, University of Mount Union, USA.

*Correspondence: Saundenw@mountunion.edu

Abstract

Background: The Senior Fitness Test is a reliable and valid battery of individual tests which collectively measure the strength, balance, and endurance necessary to perform activities of daily living. These tests require little equipment or cost to administer, and are commonly administered in both clinical and community settings. However, with recent and rapid advancements in technology, there is now frequent use of wearable sensors to assess and quantify mobility tasks. It was the primary aim of the present study to compare the relative and absolute test-retest reliability of body-worn sensors vs. manual assessment of Senior Fitness Test components.

Methods: Participants (65 years and older) attended a single 2-hour test session at a university academic building, where they performed two blocks of 3 tests (30-Second Chair Stand, 8-Foot Up-and-Go, and 6-Minute Walk Test) with a 45 minute to one hour rest in between blocks. Ambulatory Parkinson’s Disease Monitoring (APDM) Opal sensors and stopwatch methods simultaneously measured the intended outcomes for each test. Test-retest relative reliability for the fitness tests was assessed with intra-class correlation coefficients (ICC(2,1)). The agreement between test and retest was assessed by calculating minimum detectable change, which reflects the smallest amount of absolute change from test to retest that can be considered real.

Results: Participants included 17 females and 11 males (Age=70.6±5.4 years; Body Mass=70.6±15.3 kg; Height=1.66±0.08m). All manual 6-Minute Walk Test outcomes exhibited excellent test-retest reliability (ICC≥0.96). The same was true of APDM, except for turn duration (ICC>0.74) and the number of steps in turn (ICC>0.61). Test-retest reliability for the 30-Second Chair Stand and 8-Foot Up-and-Go were excellent for the manual method (ICC>0.83 and >0.88, respectively), but were poor for APDM (ICC>0.41 and >0.28, respectively). Minimum detectable change was similar between the APDM and manual methods for 6-Minute Walk Test cadence (6.8 vs. 6.6 steps/min, respectively) and gait speed (0.11 vs. 0.13 m/s, respectively). There were substantial differences between manual and APDM minimum detectable change for the 30-Second Chair Stand and 8-Foot Up-and-Go, with APDM requiring approximately twice as much change, compared with manual.

Conclusions: Both the APDM system and manual methods reliably and validly assessed gait speed and cadence during the 6-Minute Walk Test, though it should be noted that APDM only successfully recorded and reported data for 2 out of every 3 participants. The manual method of assessing the 8-Foot Up-and-Go and 30-Second Chair Stand was more valid and reliable than the APDM method.

Keywords: Body-worn sensors, accelerometer, inertial measurement unit, chair stand, timed up-and-go, 6-minute walk test, gait assessment, older adults

© 2017 Saunders et al; licensee Herbert Publications Ltd. This is an Open Access article distributed under the terms of Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0). This permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Introduction

There is an increasing urgency in the healthcare community to prepare for an aging population. Within the next 50 years, the number of older adults (age 65 and older) in the United States is projected to more than double in size, growing from 46 million in 2016 to over 98 million in 2060 [1]. This will increase the percentage of the older adult population in the United States from 15 percent to nearly 24 percent [1,2]. Historically, health care costs rise with advancing age related to increased number of hospitalizations, management of chronic health conditions, and dependency on long term care facilities due to decreased functional independence [3]. Furthermore, our health care system is unprepared to support the increased demands of an enlarging older adult population; there are insufficient financial resources, healthcare providers, and accessible housing facilities [3]. With changes in family structure, the rise of two-income households, and the geographic separation of extended family common in our current culture, there is frequently decreased family support available to help care for family members in ill health.

Our older adult population is growing at a rate that exceeds our ability to care for them. As the number of older adults increases, so does the call to action for the promotion of healthy aging so that older adults can remain self-sufficient and independent and maintain quality of life in their later years. Importantly, healthcare has been shifting away from reactive disease-focused care to proactive health promotion, disease prevention, and fall-risk reduction.

As discussed in the World Health Organization World Report on Ageing and Health [3], in order to understand the current patterns and characteristics of healthy aging and which interventions are the most effective, researchers and practitioners need to standardize measures for assessment of physical and functional abilities. This may help improve communication across the healthcare continuum and allow for the early identification of significant changes in performance (either positive or negative) that will improve access to services and lead to improved outcomes.

The Senior Fitness Test (SFT), developed by Rikli and Jones, is a reliable and valid battery of individual tests with published normative data [4] and criterion-referenced standards [5]. The 30-Second Chair Stand, 8-Foot Up-and-Go, and 6-Minute Walk Test are components of the SFT which collectively measure the strength, balance, and endurance necessary to perform activities of daily living [4]. These tests require little equipment or cost to administer, and are commonly administered in both clinical and community settings. However, the outcomes are limited. For example, the only outcome for 6-Minute Walk Test is the total distance traveled in six minutes. It may be useful to know cadence, steady-state gait speed, and turn durations, as these parameters may change even if total distance does not.

With recent and rapid advancements in technology, there is now frequent use of wearable sensors to assess and quantify mobility tasks [6]. Ambulatory Parkinson’s Disease Monitoring (APDM) is one such system of body-worn sensors that has the potential to improve the reliability and validity of SFT outcomes, and it permits the automated measurement of parameters that would otherwise be difficult to measure with the human eye (e.g., step duration, trunk lean angle, turning velocity, sit-to-stand duration, etc.). However, with these possible advantages comes the responsibility of evaluating its usability and error.

In addition to the standard practice of reporting statistical significance and effect sizes, it is suggested that researchers more commonly present minimum detectable change (MDC) [7], the smallest magnitude of change that can be considered real change, rather than chance or error. While the MDC for manual assessment of the 30-Second Chair Stand [8-10], 8-Foot Up-and-Go [8-10] (or various versions of the Timed Up-and-Go [11-13]), and the 6-Minute Walk Test [10-13] have been established, to our knowledge the MDC for instrumented (APDM) and manual versions of these tests have not been directly compared. Furthermore, the MDC for cadence, gait speed, and turn duration during the 6-Minute Walk Test have yet to be established. It was therefore the primary aim of the present study to compare the relative and absolute test-retest reliability of the APDM vs. manual assessment of the 30-Second Chair Stand, 8-Foot Up-and-Go, and 6-Minute Walk Test. A secondary aim was to evaluate the concurrent validity of APDM and manual assessment of the same tests. It was hypothesized that APDM and manual assessments would demonstrate excellent correlation and agreement, but that APDM would minimize human error, which would result in smaller MDC values, compared with manual test outcomes.

Methods

Participants

Participants were older adults (65 years and older) who were able to walk unassisted, had no diagnosed neuromuscular or cognitive disorders, and had not experienced a fall within the last 12 months. This study was approved by the University of Mount Union Institutional Review Board, and all participants signed an informed consent prior to any testing.

Study design

The primary purpose of this study was to establish MDC values for the 30-Second Chair Stand, 8-Foot Up-and-Go, and 6-Minute Walk Test, when assessed by body-worn inertial measurement units and manual recording with a stopwatch. Participants attended a single 2-hour test session at a university academic building, where they performed two blocks of all 3 tests with a 45 minute to one hour rest in between. Though no rigid rest protocol was utilized, all participants were asked to remain seated during all rest periods. Sensor and stopwatch methods simultaneously measured the intended outcomes for each test. While each participant was present for two hours, only about 20 total minutes of physical activity was performed. Given the ample rest between tests and blocks of tests, fatigue was not expected to impact test performance.
Measurement and Instrumentation

Ambulatory Parkinson’s Disease Monitoring (APDM) Opal sensors are body-worn, battery operated, wireless sensors that were developed by APDM Wearable Technologies (Portland, OR). Each of the six Opal sensors consists of a tri-axial accelerometer, gyroscope, and magnetometer. The wireless transmission range for the sensors is 30 m line of sight. The sensors were affixed to the posterior aspect of both wrists, dorsal side of the feet (approximately atop the talus), the sternum, and lumbar region on the back using Velcro straps. A Bluetooth remote that came with the system was used to start all tests and stop only the 8-foot Up-and-Go. Upon completion of each test, the data was immediately transmitted wirelessly to a Windows laptop running Mobility Lab software, also created by APDM Wearable Technologies. The sensors, in conjunction with Mobility Lab software, are able to immediately deliver spatial and temporal outcomes for the 30-Second Chair Stand, 8-Foot Up-and-Go, and 6-Minute Walk Test.

The manual method of assessing the fitness tests required only a generic cell phone stopwatch application capable of recording and storing lap times.

Procedures

Participants self-reported their age and weight (weight was not measured, as it had little significance to this investigation). Height was measured with a wall mounted tape measure. Given that no cross-sectional age or gender comparisons were being made, we elected to measure height with shoes on to improve the feasibility of the test session. Two blocks of tests occurred approximately one hour apart. In each block, participants performed (in order) a 2-minute familiarization trial of the 6-Minute Walk Test, 5-minute rest, familiarization trial of the 30-Second Chair Stand, 2-minute rest, recorded trial of the 30-Second Chair Stand, 2-minute rest, familiarization trial of the 8-Foot Up-and-Go, 2-minute rest, recorded trial of the 8-Foot Up-and-Go, and finally the 6-Minute Walk Test. Prior to each block of tests, APDM opal sensors were affixed on both wrists, tops of the feet (approximately atop the talus), the sternum, and lumbar region. The components of the Senior Fitness Test employed in this study (30-Second Chair Stand, 8-Foot Up-and-Go, and 6-Minute Walk Test) are described in greater detail elsewhere [4]. Briefly:

30-Second chair stand

Participants began in a seated position in an approximately 17-inch-high armless chair positioned with its back against a wall for stability. Each participant was instructed to cross their arms over their chest, maintain an approximately shoulder-width stance, and completely stand and sit as many times as possible in 30 seconds. If participants felt they could not complete the test with their arms over their chest, they were given the option to place their hands on their knees. Any corrections to technique were communicated to the participant during the familiarization trial. On the word “go,” one investigator initiated the APDM recording, and another investigator counted chair stands. Following one familiarization trial, a single experimental trial was recorded.

8-Foot Up-and-Go

The same chair and chair position was used for this test. Participants were instructed to begin seated with their feet in a slightly staggered stance and hands on knees. They were then directed to stand up from the chair, walk as fast as they comfortably could, without running, around an object placed on the ground 8 feet from the front of the chair, and return to their initial seated position in the chair. Any corrections to technique were communicated to the participant during the familiarization trial. On the word “go,” one investigator initiated the APDM recording, and another investigator simultaneously started their stopwatch. Both investigators stopped their respective devices when the participant returned to a seated position. Following one familiarization trial, a single experimental trial was recorded.

6-Minute Walk Test

The 6-Minute Walk Test was completed on a 25 m, linear course (Figure 1). Ends of the course were marked with a line of blue masking tape (lines A and D), and dashed lines

Figure 1. Methodology for the manual assessment of stride rate and steady-state gait speed during a 6-Minute Walk Test. Points A and D represent 180° turns, while the area between points B and C represent a 20 meter area where gait speed was assumed to be relatively constant. Numbers 1-10 indicate the counting of strides. Start and Lap indicate functions on a stopwatch.
were positioned 2.5 meters from each end (lines B and C). Participants were instructed to completely pass each end line before turning, and to use the turning style of their own choosing. It was communicated that the goal was to walk at the fastest pace they felt they could maintain for six minutes. Chairs were positioned at each end of the course in case a break was necessary. The participant’s position on the course at the termination of 6 minutes was marked for measurement of the partial distance traveled in the final lap.

On the word “go,” one investigator initiated the APDM recording, a second investigator began the manual assessment of gait speed and turn duration, and a third investigator began the manual assessment of cadence.

Gait speed was manually assessed with a stopwatch following these instructions (Figure 1):
- Press “Start” as the participant crosses point B
- Press “Lap” every time the participant crosses points B and C
- Press “Stop” when the trial is completed, regardless of where they are on the course
- Enter all lap times (to 2 decimal places) for all completed interior 20m segments (i.e., longest times in your list)
- Enter all lap times (to 2 decimal places) for all turns (i.e., shortest times in your list)

Knowing the time to complete each 20 m distance, steady-state gait speed was calculated as 20 divided by the average 20 m lap time. Overall gait speed was calculated as the total distance walked in 6 minutes divided by 360 seconds. This permitted us to differentiate between steady-state gait speed and the gait speed including decelerations and accelerations into and out of turns. Turn duration was simply calculated as the average time taken to go from the end steady-state marker, turn, and return to the beginning steady state marker for the new lap. Note that the manual assessment of turn duration included decelerating and accelerating gait into and out of turns.

Stride rate was manually assessed with a stopwatch following these instructions (Figure 1):
- Press “Start” as the participant’s right foot touches the ground past point B
- Starting with the next RIGHT step, begin counting right steps 1-10
- Press “Lap” on the 10th RIGHT step
- Press “Lap” as the participant’s right foot touches the ground nearest point C on their return
- Starting with the next RIGHT step, begin counting right steps 1-10
- Press “Lap” on the 10th RIGHT step
- Press “Stop” when the trial is completed, regardless of where they are on the course
- Enter all lap times (to 2 decimal places) for all completed 10-stride segments

Cadence was calculated by converting the average time to complete 10 strides to the number of steps per minute (considering 2 steps/stride and 60 s/min).

Senior Fitness Test components not presently capable of being assessed by APDM, and therefore not included in the present study are the Arm Curl (the number of arm curls performed in 30 seconds), Back Scratch (distance between fingertips when reaching behind the back from above with one arm and from below with the other), Chair Sit-and-Reach (distance between fingertips and toes when reaching for the toes from a seated position in a chair), and 2-Minute Step (number of right steps taken when marching in place for 2 minutes).

**Inter-Rater Reliability**

In a deliberate attempt to improve the external validity of this study, 10 investigators with different levels of education and experience (one Physical Therapy professor, one Exercise Science professor, and eight undergraduate Exercise Science students) conducted the manual assessment of the Senior Fitness Test components. Importantly, the same investigator assessed the test and retest for a given participant. While the inter-rater reliability of the 30-Second Chair Stand, 8-Foot Up-and-Go, and 6-Minute Walk Test distance are well established as being excellent, to our knowledge, we are the first to present methodology for manually assessing steady-state gait speed and cadence during the 6-Minute Walk Test. As such, it was important to establish the inter-rater reliability for these outcomes prior to comparing them with APDM. A video recording was made of a single participant completing the 6-Minute Walk Test. Each of 10 raters, using the methods described above, watched the video one time to assess gait speed, and watched it once more to assess cadence. Inter-rater reliability was then assessed for each gait parameter.

**Statistical analyses**

Statistical analyses were conducted using SPSS version 24 (IBM Corp, Armonk, NY). Inter-rater reliability for the manual recording of 6-Minute Walk Test mean steady state gait speed and cadence was assessed with intra-class correlation coefficients (ICC1,2). Test-retest relative reliability for SFT components and subcomponents (APDM and manual) was also assessed with ICC1,2. Because ICCs are very much influenced by the heterogeneity of the population being studied [14], it should be noted that the results of these analyses may only be reasonably applied to individuals of similar age and functional ability.

The agreement between test and retest, which is more stable over population samples [15], was determined with two methods. First, MDC$_{95}$ was calculated as:

$$1.96 \times \sqrt{2} \times \text{SEM}$$

where $\text{SEM} = \text{SD} \times \sqrt{1 - \text{ICC}}$ and $\text{SD} = \sqrt{\frac{\text{S}_{\text{Total}}}{n - 1}}$ from ANOVA.

MDC$_{95}$%, the percent change in performance required to be minimally detectable, was calculated as:

$$\frac{\text{MDC}}{\text{baseline mean}} \times 100.$$
significant difference between test and retest.

Bland-Altman plots (with 95% Limits of Agreement) and paired t-tests were used to test for differences between APDM and manual counting or stopwatch timing (i.e., concurrent validity) for each fitness test outcome. Furthermore, regression analyses were performed to look for proportional bias in the relationship between the mean of two measures and the difference between them. When proportional bias was present, the ratio of the two measures was plotted against the mean of the two measures, such that a correction factor could be calculated to account for systematic bias regardless of the magnitude of the outcome [16]. MDC95% (from the test-retest analyses) was also useful in comparing the sensitivity of APDM vs. manual recording, with the smaller of the two values representing the more sensitive recording technique. This is especially useful when two techniques may not be measuring the same thing and an absolute comparison wouldn’t make sense (e.g. turn duration for the 6-Minute Walk Test). Data from blocks 1 and 2 (i.e., before and after the 1-hour rest) were collapsed for the measurement method comparison. Importantly, the APDM sensors failed to record trials for 9 of the 28 participants. Only the 19 participants with both APDM and manual data were included in the concurrent validity analysis. Significance for all statistical analyses was established a priori at alpha = 0.05.

Results

Baseline Subject Characteristics
Participants included 17 females and 11 males (Age=70.6±5.4 years; Body Mass =70.6±15.3 kg; Height =1.66±0.08m).

Inter-Rater Reliability
Figure 2 shows the cadence and steady-state gait speed for each of 19 laps during a 6-Minute Walk Test by a single participant. The values presented reflect the mean and standard deviation of 10 raters assessing the same test on video. Note that both the cadence and steady-state gait speed fluctuated during the 6-minute test, and that all 10 raters were sensitive to that change. The mean cadence and steady-state gait speed for the entire 6-Minute Walk Test from each of the raters exhibited excellent inter-rater reliability (ICC>0.92 and 0.98, respectively).

Test-Retest Reliability
Table 1 presents the test-retest reliability and agreement for the 6-Minute Walk Test, 30-Second Chair Stand, and 8-Foot Up-and-Go, using APDM and manual measurement methods. A direct comparison between APDM and manual methods was made for the 19 participants with a complete dataset from each (normal font), while the italicized values represent the manual method data in isolation for all 28 participants. It should be noted that the manual method outcomes from all 28 participants are only marginally improved from the manual method outcomes for the 19 participants with matched APDM data.

All manual 6-Minute Walk Test outcomes (cadence, gait speed, steady-state gait speed, distance, and turn duration) exhibited excellent test-retest reliability (ICC≥0.96). The same
was true of APDM, except for turn duration (ICC > 0.74) and the number of steps in turn (ICC > 0.61). Test-retest reliability for the 30-Second Chair Stand and 8-Foot Up-and-Go were excellent for the manual method (ICC > 0.83 and > 0.88, respectively), but were poor for APDM (ICC > 0.41 and > 0.28, respectively).

While most 6-Minute Walk Test test and retest outcomes were highly correlated (for both APDM and manual methods), paired t-tests revealed significantly improved performance in the second trial for all outcomes, except for APDM turn duration and number of steps in turn. No significant trial 1 to trial 2 changes were indicated for the 30-Second Chair Stand or 8-Foot Up-and-Go, for either method.

MDC, which reflects the smallest amount of absolute change from test to retest that can be considered real, was similar between the APDM and manual methods for 6-Minute Walk Test cadence (6.8 vs. 6.6 steps/min, respectively) and gait speed (0.11 vs. 0.13 m/s, respectively). The only other 6-Minute Walk Test outcome assessed by both APDM and manual methods was turn duration, though it should be noted that APDM and the manual method did not measure the same thing. APDM was measuring the actual angular accelerations that accompany the initiation and termination of a turn. The manual assessment included the decelerating footsteps preceding the turn and the accelerating footsteps following the turn. Therefore, it makes more sense to compare the relative change in performance (MDC%), rather than the absolute change (MDC). The manual method required less percent change from baseline than APDM to be considered real (12.6% vs. 20.7%). There were also substantial differences between manual and APDM MDC for the 30-Second Chair Stand and 8-Foot Up-and-Go, with APDM requiring approximately twice as much change, compared with manual.

There were a few 6-Minute Walk Test outcomes that were only measured by APDM or the manual method, but not both. Notably, the manual assessment of steady-state gait speed and total distance, and the APDM assessment of step duration, resulted in MDC% < 10%. The APDM derived number of steps in turn produced a much larger MDC% (28.4%).

### Concurrent Validity

Bland-Altman Plots were created to show the level of agreement between the APDM and manual methods for the assessment of SFT outcomes (Figures 3-5). The solid horizontal lines reflect the average difference between the measurement techniques (bias) and should ideally be zero. The dashed horizontal lines represent ± 95% limits of agreement (random error), and should be as narrow as possible.

APDM and manual methods agreed well for cadence (Figure 3A) and gait speed (Figure 3B), as indicated by almost zero bias (supported by non-significant paired t-tests) and reasonably small random error components (±2.72 steps/min for cadence, and ±0.077 m/s for gait speed). There was, however, systematic bias indicated when comparing APDM gait speed to the manually assessed steady-state gait speed (significant paired t-test), with steady-state gait speed exceeding APDM gait speed by an average of 0.1 m/s (Figure 3C). Regression analysis also indicated proportional bias (i.e., the between-method difference in the reported gait speed increased with increasing mean gait speed). To account for the proportional bias, the ratio of the steady-state to APDM gait speed was plotted against the mean of the two measures (Figure 3D). On average, the steady-state gait speed was 1.066 times faster than the APDM gait speed. After applying this correction factor to the individual APDM gait speeds (Figure 3E), there were no systematic bias (paired t-test no longer significant), and the random error was narrow (±0.045 m/s).

On average, APDM reported 0.5 fewer chair stands than the manual method (Figure 4), a rather marginal but significant bias, but the random error was large (±2.5 chair stands). Unlike the 30-Second Chair Stand, for the 8-Foot Up-and-Go

---

**Table 1. Test-retest reliability and agreement of Senior Fitness Test components.**

| Outcomes                  | ICC     | MDC     | MDC (%) |
|---------------------------|---------|---------|---------|
|                           | APDM    | Manual  | APDM    | Manual  | APDM    | Manual  |
|                           | (n = 19) | (n = 28) | (n = 19) | (n = 28) | (n = 19) | (n = 28) |
| Cadence (steps/min)       | 0.96    | 0.96    | 0.68    | 0.66    | 0.79    | 0.70    |
| Gait Speed (m/s)          | 0.97    | 0.97    | 0.11    | 0.13    | 0.10    | 0.12    |
| Steady-State Gait Speed (m/s) | 0.94    | 0.94    | 0.99    | 0.99    | 0.94    | 0.99    |
| Total Distance (m)        | 0.98    | 0.98    | 0.03    | 0.03    | 0.46    | 0.37    |
| Step Duration (s)         | 0.74    | 0.97    | 0.42    | 0.65    | 0.24    | 0.24    |
| Turn Duration (s)         | 0.61    | 0.97    | 1.1     | 1.1     | 0.46    | 0.37    |
| Steps in Turn (%)         |         |         | 0.61    | 0.97    | 0.46    | 0.37    |
| Chair Stand (s)           | 0.41    | 0.83    | 2.7     | 2.6     | 51      | 27      |
| 8 Foot Up-and-Go Time (s) | 0.28    | 0.88    | 1.02    | 1.02    | 48.8    | 21.8    |

Note. ICC=Intraclass Correlation Coefficient; MDC=Minimum Detectable Change; MDC (%)=MDC expressed as the percent change from the baseline group mean; APDM=body-worn sensors.
Figure 3. Bland-Altman Plots illustrating the agreement between APDM and manual methods for 6-Minute Walk Test Outcomes. The solid horizontal lines represent the average difference between APDM and manual methods (systematic bias), while the dashed lines represent the ±95% Limits of Agreement (random error).

(Figure 5), APDM reported times that were on average 1.35 s longer than the manual method (significant paired t-test), and exhibited a random error similar in magnitude (±1.23 s). Because of the great magnitude of the random error, a correction factor was not applied here.

Discussion
As new technology is developed to enhance the assessment of gait and other functional movements, it is imperative that the reliability and validity is well established for diverse populations. This study aimed to evaluate the concurrent validity and test-retest MDC of the APDM system for the assessment of three SFT components. Contrary to our hypothesis, the manual method produced MDCs that were as small as or smaller than APDM for all gait outcomes and about 50% smaller than APDM for the 30-Second Chair Stand and 8-Foot Up-and-Go. In support of our hypothesis, there was excellent agreement between APDM and the manual method for overall gait speed and cadence during the 6-Minute Walk Test. However, a correction factor needed to be applied to the
APDM gait speed for it to agree well with the manual method steady-state gait speed. Additionally, systematic bias was evident for the 30-Second Chair Stand and 8-Foot Up-and-Go, with APDM underreporting the number of chair stands and over reporting the time taken to complete the 8-Foot Up-and-Go, compared with the manual method.

**Test-Retest Reliability**

The MDC for the same test has been shown to vary substantially between populations, with clinical populations demonstrating greater MDCs, compared with healthy controls [11,12,17-19]. Given that much of what we know regarding the MDC for the tests utilized in the present study were reported for populations with diagnosed pathologies, it makes direct comparisons between studies difficult. Still, the MDCs presented here for the manual method of assessing 6-Minute Walk Test distance (37.7-46.6 m) are strikingly similar to the range of MDCs reported previously (39.2 m [11], 44.1 m [10], and 82 m [13]), despite differences in population, course design, and rest intervals. Likewise, the MDCs for the manual timing of the 8-Foot Up-and-Go in the present study (1.02-1.47 s) are well within the range of values already published for the 8-Foot Up-and-Go (1.19-1.70 s [8-10]) and related 3-meter Timed Up-and-Go (1.1-1.50 s [11,13,20]). The present manual 30-Second Chair Stand MDCs (2.6-2.7 stands) are on par with the 2.4 stands reported for patients with mild cognitive impairment [10], but larger than those reported for community-dwelling older adults (1.2-1.3 stands [8,9]). Differences, while noteworthy, are small enough in magnitude that they can easily be attributed to any differences in age, fitness level, and motivation of the populations studied.

To our knowledge, we are the first to report MDCs for these SFT components using the APDM system. And while the MDCs for the 6-Minute Walk Test are similar to the manual method here and elsewhere, the MDCs for the 8-Foot Up-and-Go and 30-Second Chair Stand are unacceptably high, to the extent that they likely well exceed expected improvements following clinical treatment or fitness training. For example, a person completing 10 chairs stands at baseline and 13 chair stands at follow-up would indicate significant improvement when assessed manually, but not if assessed by APDM. The same is true for the 8-Foot Up-and-Go.

The poor reliability of the APDM for these tests is likely a result of the proprietary algorithm created to calculate the outcomes. Though not disclosed to the consumer, it is likely the angular acceleration of the sternum and lumbar sensors that detect the initiation and completion of a chair stand. It is reasonable to assume that variability between the two trials in number of completed chair stands is related to the variability in mechanics (e.g., trunk lean angles), rather than the actual number of stands.

Unlike the 30-Second Chair Stand, there was clearly extensive random error associated with the timing of the 8-Foot Up-and-Go, which is rather odd given that it is timed in an identical manner to a stopwatch (i.e., the investigator starts and stops the timer by manually clicking a Bluetooth remote). This should have agreed well with the manual method. If anything, the error should have been almost entirely systematic, with APDM consistently being fast or slow by the same magnitude. The time reported by APDM was longer than manual timing for every trial, but it was not consistent enough to warrant
the application of a correction factor.

**Concurrent Validity**

The findings here suggest good agreement between the APDM and manual methods for average cadence and gait speed, but not steady-state gait speed, for the 6-Minute Walk Test. This is not a surprising finding since APDM calculates gait speed step-by-step and is most certainly including accelerating and decelerating steps in the calculations, which results in a consistently slower gait speed than the steady-state gait speed. Hollman et al. [21] found no significant difference between APDM gait parameters for a 6-Minute Walk Test performed on a treadmill or over the ground. This might appear to contradict our findings, as treadmill walking represents a steady-state gait condition and over the ground included many 180° turns on a 42 meter linear course. However, they filtered out turns and decelerating gait cycles prior to making the comparison, a process not employed here. We derived a correction factor that effectively compensates for the systematic difference between the gait speed reported by APDM and actual steady-state gait speed. Future users of the APDM system should multiply the reported gait speed by 1.066 to get a good approximation of the steady-state gait speed. It should be noted that this correction factor specifically applies to our 6-Minute Walk Test protocol using a 25 m course, and that any alteration to the protocol, especially altering the length of the course, is likely to change the relationship between APDM gait speed and steady-state gait speed.

Though turn duration using the APDM and manual methods should not be directly compared (because the two methods are measuring different things), the difference in test-retest reliability between the two methods is worthy of discussion. Test-retest reliability for the manual method of assessing turn duration (ICC > 0.98) suggests participants were quite similar between trials with respect to how they turn. The much lower ICC for APDM turn duration (ICC > 0.74) indicates greater variability in the components used to calculate turn duration than the actual turn duration. This would also explain the poor test-retest reliability for the number of steps in turn, as steps contributing to a turn would only be counted if they occur in the interval between when APDM senses the initiation and termination of a turn. It might be prudent for APDM to reevaluate the filter used to detect turn initiation and termination.

As with turn duration during the 6-Minute Walk Test, APDM and manual 8-Foot Up-and-Go times were considerably different (APDM was on average 1.35 s longer). Coulthard et al. [22] also reported that APDM correlated highly with a stopwatch for a 3.5 m Timed Up-and-Go, but overestimated time by about 2 seconds. This consistency across studies suggests there is an inherent flaw in how APDM detects the beginning and end of an Up-and-Go trial, rather than investigator error. We believe we are the first to formally compare the number of chair stands reported by APDM and manual counting. On average, APDM reported 0.5 fewer chair stands than the manual method. While statistically significant, it is unlikely that this difference would alter the prognosis of a patient. However, recall that APDM was less reliable than the manual method in a test-retest scenario, suggesting it might be better to simply count.

**Limitations**

There were several limitations to this study. First, a rather homogeneous healthy population of older adults threaten the external validity, while the variability in education and experience level of the investigators threaten the internal validity. Though care was taken to establish the inter-rater reliability of the investigators, only one video analysis was performed and intra-rater reliability was not assessed. Second, regarding the repeated design methodology, step width for the beginning position of the 8-Foot Up-and-Go was not measured, and this could have led to greater trial-to-trial variability, but note that this does not influence result interpretation when comparing the test-retest reliability of manual and APDM methods. Finally, the results of this study are specifically relevant to other healthy older adults, and care should be taken when interpreting these findings as they might relate to special populations.

Of particular note, the APDM system failed to report complete data for 33% of our participants. Additionally, APDM provides many other outcomes than those investigated here (e.g., step duration, trunk accelerations, lean angles, sit-to-stand and stand-to-sit durations, etc.), but data for those outcomes were missing too often to be included in the present study. While it is possible that this resulted from a failure of our computer or investigator error, the lack of data reporting appeared to be random rather than systematic, while our computer and protocols were constant.

**Conclusions**

For tests where APDM successfully recorded and reported data (only 2 out of every 3 tests), the APDM system proved to be a reliable and valid method of assessing gait speed and cadence during a 6-Minute Walk Test, but turn duration and number of steps in turn are too variable to be clinically useful. It is recommended that APDM reassess the detection of turn initiation and termination. A good approximation of steady-state gait speed can be calculated by multiplying the APDM gait speed by 1.066, and this correction factor was shown to be accurate across a broad range of gait speeds.

To our knowledge, we are the first to provide methodology for manually assessing steady-state gait speed, cadence, and turn duration during the 6-Minute Walk Test, and it resulted in acceptably small MDCs for those outcomes. The manual method of assessing the 8-Foot Up-and-Go and 30-Second Chair Stand was more valid and reliable than the APDM method. The APDM system is capable of providing numerous outcomes unobtainable via manual means (e.g., step...
duration, trunk accelerations, lean angles, sit-to-stand and stand-to-sit durations, etc.), and these outcomes are worthy of further evaluation to determine their reliability, validity, and clinical usefulness.

Competing interests
The authors declare that they have no competing interests.

Acknowledgements
While their contributions did not warrant authorship, we would like to thank (in no particular order) Abigail Matsushima, Uriel Ibarra-Moreno, Kennedy Miller, Morgan Kiser, and Alyssa Braun for their integral role in designing this study, collecting data, and interpreting the results. We would also like to express appreciation to Dr. Jacob Barkley, PhD for his statistical expertise.

References
1. Mathur M, Jacobsen LA and Pollard KM. Aging in the United States. Population Bulletin. 2015.
2. Otmran JM and Velkoff VA. An Ageing Nation: The Older Population in the United States. U.S. Census Bureau: Washington DC. 2014; 25-1140.
3. World Report of Ageing and Health, 2015, World Health Organization: Geneva, Switzerland.
4. Rikli RE and Jones CJ. Functional Fitness Normative Scores for Community-Residing Older Adults, Ages 60-94. J Aging Phys Act. 1999; 7:162-181. | Pdf
5. Rikli RE and Jones CJ. Development and validation of criterion-referenced clinically relevant fitness standards for maintaining physical independence in later years. Gerontologist. 2013; 53:255-67. | Article
6. Horak F, King L and Mancini M. Role of body-worn movement monitor technology for balance and gait rehabilitation. Phys Ther. 2015; 95:461-70. | Article
7. Haley SM and Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. Phys Ther. 2006; 86:735-43. | PubMed
8. Bhattacharya PK, Deka K and Roy A. A community-based study to assess test–retest reliability of senior fitness test in the geriatric population in a northeastern Indian city. Int J Med Sci Public Health. 2015; 5.
9. Bhattacharya PK, Deka K and Roy A. Assessment of inter-rater variability of the Senior Fitness Test in the geriatric population: A community based study. Int J Biomed Adv Res. 2016; 7:208-212.
10. Hesseberg K, Bentzen H and Bergland A. Reliability of the senior fitness test in community-dwelling older people with cognitive impairment. Physiother Res Int. 2015; 20:37-44. | Article
11. Quinn L, Khalil H, Dawes H, Fritz NE, Kegelmyer D, Kloos AD, Gillard JW and Busse M. Reliability and minimal detectable change of physical performance measures in individuals with pre-manifest and manifest Huntington disease. Phys Ther. 2013; 93:342-56. | Article
12. Ries JD, Echternach JL, Nof L and Gagnon Blodgett M. Test-retest reliability and minimal detectable change scores for the timed “up & go” test, the six-minute walk test, and gait speed in people with Alzheimer disease. Phys Ther. 2009; 89:569-79. | Article
13. Steffen T and Seney M. Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-item short-form health survey, and the unified Parkinson disease rating scale in people with parkinsonism. Phys Ther. 2008; 88:733-46. | Article
14. Atkinson G and Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998; 26:217-38. | PubMed
15. de Vet HCW, Terwee CB, Knol DL and Bouter LM. When to use agreement versus reliability measures. J Clin Epi. 2006; 59:1033-1039. | Article
16. Bland JM and Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999; 8:135-60. | Article
17. Hars M, Herrmann FR and Trombetti A. Reliability and minimal detectable change of gait variables in community-dwelling and hospitalized older fallers. Gait Posture. 2013; 38:1010-4. | Article
18. Peters DM, Middleton A, Donley JW, Blankel EL and Fritz SL. Concurrent validity of walking speed values calculated via the GAITRite electronic walkway and 3 meter walk test in the chronic stroke population. Physiother Theory Pract. 2014; 30:183-8. | Article
19. Spain RI, Mancini M, Horak FB and Boudette D. Body-worn sensors capture variability, but not decline, of gait and balance measures in multiple sclerosis over 18 months. Gait Posture. 2014; 39:958-64. | Article
20. Alghadir A, Anwer S and Brismee JM. The reliability and minimal detectable change of Timed Up and Go test in individuals with grade 1-3 knee osteoarthritis. BMC Musculoskelet Disord. 2015; 16:174. | Article
21. Holliman JH, Watkins MK, Imhoff AC, Braun CE, Akervik KA and Ness DK. A comparison of variability in spatiotemporal gait parameters between treadmill and overground walking conditions. Gait Posture. 2016; 43:204-9. | Article
22. Coughard JT, Treen TT, Oates AR and Lanovaz JL. Evaluation of an inertial sensor system for analysis of timed-up-and-go under dual-task demands. Gait Posture. 2015; 41:882-7. | Article

Citation:
Salvatore MD, Colacino AM, Hess ME, Todd SW and Saunders NW. Concurrent validity and minimum detectable change of Senior Fitness Test components: instrumented vs. manual assessment. Phys Ther Rehabil. 2017; 4:13.
http://dx.doi.org/10.7243/2055-2386-4-13