Title page

Reliability and validity of four step tests in older adults with dementia

Authors:

Wayne Lap Sun Chan [1]

Sharon Man Ha Tsang [1]

Lily Yuen Wah Ho [2]

Affiliation:

[1] Department of Rehabilitation Sciences, The Hong Kong Polytechnic University

[2] School of Nursing, The Hong Kong Polytechnic University

Acknowledgement:

We want to thank the participants and their main caregivers in this project and the staff at multiple community centers for older adults who helped coordinate the recruitment and arranged the venue for the assessment sessions.

The materials of this manuscript have never been presented before in any circumstance. This feasibility trial was funded by The Hong Kong Polytechnic University (Project ID: P0030116).

Conflict of interest statement:

We declare that the authors did not have any conflict of interest.

Corresponding author:

Dr. Chan, Wayne L. S.

Department of Rehabilitation Sciences,

ST 505, 5/F, Ng Wing Hong Building,

The Hong Kong Polytechnic University

Email: wayne.ls.chan@polyu.edu.hk

Office number: +852 2766 6742

Clinical Trial Registration number:

NCT04296123

Abstract

Objectives

To determine the test-retest and inter-rater reliability, concurrent and discriminative validity of the Four Square Step Test (FSST), the Choice Stepping Reaction Time Test (CSRTT), the Maximum Step Length Test (MSLT), and the Alternate Step Test (AST) in older adults with dementia.

Methods

Thirty-seven older adults with dementia who could walk independently for at least 10 m were recruited at community centers and day care centers for older adults. The participants completed the step tests conducted by two independent raters on three separate testing occasions within three weeks. In addition, the physical and cognitive function of the participants were evaluated at baseline.

Results

The FSST, CSRTT, and MSLT showed good-to-excellent test-retest reliability (intraclass correlation coefficient (ICC) = 0.83–0.91), and the AST exhibited fair test-retest reliability (ICC = 0.70). All the step tests showed good-to-excellent inter-rater reliability (ICC = 0.75–0.94). The step tests had moderate-to-strong correlations with various physical and cognitive measures (Pearson correlation coefficients = 0.34–0.72). The MSLT side step and AST could differentiate between individuals who did and did not use a walking stick to ambulate ($p \leq .046$). The FSST, CSRTT, and AST could differentiate between individuals with and without a major neurocognitive impairment ($p \leq .005$).

Conclusion

The FSST, CSRTT, and MSLT were reliable and valid for examining the stepping performance of older adults with dementia. Clinicians can use these tests to evaluate the

physical and cognitive function of this population and identify those with significant cognitive impairment.

Keywords

Dementia; Clinical assessment; Falls; Stepping performance

Key points

- The Four Square Step Test, the Choice Stepping Reaction Time Test, and the Maximum Step Length Test had good-to-excellent test-retest and inter-rater reliability. The Alternate Step Test had only modest test-retest reliability and good inter-rater reliability.
- All the four step tests had moderate-to-high correlations with different physical and cognitive outcome measures, indicating that they had generally good-to-excellent concurrent validity.
- The Maximum Step Length Test and the Alternate Step Test could differentiate between people with dementia with and without using a walking stick to ambulate. The Four Square Step Test, the Choice Stepping Reaction Time Test, and the Alternate Step Test could differentiate between people with dementia with and without a major neurocognitive impairment.
- Clinicians can use the Four Square Step Test, the Choice Stepping Reaction Time Test, and the Maximum Step Length Test to evaluate the physical and cognitive function of people with dementia and identify those with significant cognitive impairment.

Introduction

The ability to walk is a significant indicator of overall health and well-being in older adults. Walking ability predicts functional independence, hospital admission, and mortality in older adults [1,2]. Poor walking ability is also a risk factor for dementia in older adults [3,4]. Furthermore, the walking ability of older adults with dementia is reduced [5]. Gait impairments (e.g., reduced walking speed, reduced step length, and increased gait variability) are associated with falls, functional dependence, and mortality in older adults with dementia [6–10]. Therefore, studying gait-relevant parameters and ambulatory components associated with walking ability can be used to predict health outcomes in this population.

Stepping is a daily task that older adults need to perform to walk, maintain their posture and balance. When older adults walk in the community and encounter environmental challenges (e.g., uneven ground and slippery surfaces) and external perturbations (e.g., (e.g., trip over an obstacle or are bumped by other people), they must adjust their steps (e.g., changing their step length, width, and height) or increase their base of support through stepping to regain balance and prevent falls [11–13]. Taking fast and precise steps is crucial for older adults to walk safely, maintain balance, and prevent falls. Although walk tests have been widely used to evaluate the walking ability of older adults with dementia [14] and may partially evaluate their stepping performance, walk tests usually assume individuals to walk in a controlled environment (e.g., on a leveled ground and a straight corridor in one direction). Therefore, developing specific tests to evaluate stepping performance in demanding circumstances (e.g., taking bigger steps to overcome obstacles, taking steps in changing directions, and avoiding wrong step placements) is essential for clinicians to determine the levels of independence of older adults when ambulating in the community.

The stepping performance of older adults is determined by their sensorimotor (e.g., reaction time and muscle power) and cognitive function (e.g., attentional control and visuospatial orientation) [15]. Older adults with dementia often have significantly deteriorated cognitive function, particularly executive function, which decreases their ability to integrate

sensory information and implement appropriate motor responses. In addition, older adults with dementia often have a reduced ability to mobilize their cognitive resources in balance-challenging conditions to maintain their balance and thus have an increased risk of falls [16]. Previous studies have reported that falls are common in older adults with dementia, and impaired stepping (e.g., reduced step length and increased step width during walking) is one of the factors leading to falls in this population [6,17]. Thus, a reliable and valid tool for evaluating the stepping performance of older adults with dementia would help assess balance and predict fall risk in this population.

Several step tests, such as the Four Square Step Test (FSST) [18], the Choice Stepping Reaction Time Test (CSRTT) [19], the Maximum Step Length Test (MSLT) [20], and the Alternate Step Test (AST) [21], have been developed to evaluate the stepping performance of older adults. The FSST assesses the ability of older adults to execute a predetermined step sequence that involves stepping in four different directions (e.g., forward, backward, and sideways). The CSRTT evaluates how fast older adults respond to the verbal instructions of an assessor to take steps correctly. The MSLT measures the maximum step length that older adults can achieve without losing balance. The AST assesses the ability of older adults to shift their body weight when taking steps. These step tests have been used to evaluate the effects of different interventions in improving the stepping performance of healthy older adults [22–25]. Previous studies have shown that the aforementioned step tests are reliable and valid for assessing stepping performance and can estimate the risk of falls in healthy older adults [18–21] and other special populations [26–29]. In addition, the performance of older adults in these step tests has been found to be highly correlated with different aspects of their physical function, such as muscle strength and balance [19,30–32]. However, only one study with a small sample size (n = 14) has investigated the reliability of the Step Test developed by Hill and his colleagues in older adults with dementia [33]. The reliability of other common step tests in this population has not been investigated. Furthermore, no study was conducted to evaluate the validity of step tests in older adults with dementia. The association

between stepping performance and cognitive function in this population has yet to be thoroughly investigated.

This study aimed to determine the test-retest and inter-rater reliability and concurrent and discriminative validity of the FSST, CSRTT, MSLT, and AST in older adults with dementia. In addition, the minimal detectable changes (MDCs) of these step tests were also determined. We hypothesized that these step tests would exhibit good test–retest and inter-rater reliability, be moderately correlated with other physical and cognitive outcome measures, and could differentiate older adults with dementia with different levels of ambulatory abilities and cognitive impairment.

Materials & Methods

Participants

The participants were recruited from community centers and day care centers for older adults in Hong Kong from September 2019 to July 2023. The inclusion criteria were (i) age ≥65 years; (ii) the ability to walk independently without any walking aid or with a walking stick for at least 10 m; (iii) a physician-based, clinically-defined diagnosis of dementia or a score on the Hong Kong version of the Montreal Cognitive Assessment (HK-MoCA) of less than the age- and education-adjusted cut-offs [34]; and (iv) the presence of a main caregiver to provide informed consent. The exclusion criteria were (i) a severe musculoskeletal, cardiopulmonary, or neurological condition that restricted participation in the assessment; (ii) a substantial visual or hearing impairment that limited the ability to follow instructions; or (iii) hospitalization in the past 30 days.

Procedures

This study was conducted following the Declaration of Helsinki [35] and approved by the XXXXXXXXXXXXXXXX and the participating centers. Eligible participants and their main

caregivers were given written information and face-to-face explanations about the study. Before data collection, all participants and their main caregivers provided written consent.

An initial assessment was conducted by rater A to collect the demographic data and assess the physical and cognitive function of the participants. Personal characteristics (age, sex, education, body mass index, past medical history, and history of falls in the past six months) were collected from the reports of the main caregivers of the participants and the medical records available in the participating centers. The ambulatory status of the participants was determined based on whether they needed to use a walking stick when they walked for 10 m. The physical function of the participants was evaluated using (i) the 2-minute Walk Test (2MWT), which evaluates general mobility [14]; (ii) the 10-meter Walk Test (10mWT), which evaluates gait speed [14]; (iii) the 30-second Sit-to-Stand Test (30STS), which evaluates lower limb muscle strength [36,37]; (iv) the Berg Balance Scale (BBS), which evaluates balance [38,39]; and (v) the Chinese Disability Assessment for Dementia (CDAD), which evaluates daily functioning [40,41]. The cognitive function of the participants was examined using the HK-MoCA [34]. The participants with HK-MoCA scores equal to or less than the second percentile of the age- and education-adjusted cut-offs were regarded as having a major neurocognitive impairment [34].

The participants completed the step tests on three different testing occasions within three weeks. Two testing occasions were overseen by rater A, and one was conducted by rater B. The two raters were randomly assigned to the three testing occasions for each participant. The three testing occasions were at least one day apart from each other to reduce learning effects on the participants [42].

Two research assistants with a psychology background were appointed to be the raters. They received a 3-hour training on conducting the assessment and step tests, communicating effectively with older adults with dementia, and ensuring safety. The principal investigator (WC) trained the raters and asked them to adhere to the published guidelines when administering the step tests to the participants [42,43]. For example, the raters were

advised to use one-step commands and a calm and pleasant voice and to maintain constant eye contact with the participants when instructing them to perform the step tests. Furthermore, the raters were asked to deliver verbal and physical cues to help the participants complete the physical and step tests when necessary [42,44]. They provided such cues if a participant failed to follow the instructions and testing procedures (e.g., failed to begin, suddenly stopped moving, or walked away from the testing location).

Step tests

1. FSST [18]

The participants were asked to stand at the lower-left corner of a cross formed by four 90-cm-long walking sticks, unsupported or with a walking stick (Supplementary Figure 1a). Next, they were asked to step clockwise into each corner and retrace their steps in an anticlockwise direction as fast as possible. The time taken to finish stepping in both directions was recorded. The participants completed two stepping trials, and the average time in two trials was calculated and recorded. The FSST was previously found to be reliable (intraclass correlation coefficient (ICC) = 0.98–0.99), strongly correlated with other balance tests (Pearson correlation coefficient (*r*) = 0.47–0.88), and able to identify fallers among older adults [18].

2. CSRTT [19]

A "low-tech," non-electronic version of the CSRTT was used[45]. First, the participants were asked to stand on the standing panels on a non-slip plastic mat with both feet unsupported (Supplementary Figure 1b). Next, they were asked to step onto specific stepping panels using their right or left foot, according to a pre-determined sequence of 20 steps, as fast as possible. The first eight steps were regarded as practice trials, and the time taken to perform the final 12 steps was recorded. This version of the CSRTT was previously found to show moderate-to-good test-retest reliability (ICC = 0.74) and to be able to predict falls in healthy older adults [19].

3. MSLT [20]

The participants were instructed to stand behind the starting line marked with colored tape and fold their arms across their chests (Supplementary Figure 1c). Next, they were asked to step with their left or right leg in a specific direction (either forward, sideways, or backward) as far as possible and then return to their initial position in one step. One practice trial and three stepping trials were completed for each leg and each direction. The average step lengths of the three trials for each direction were calculated and recorded. The MSLT previously exhibited excellent test-retest reliability (ICC = 0.90–0.96) and strong correlations with balance measures (r = 0.65–0.68) in older adults [32].

4. AST [21]

This test is a modified version of one of the items in the BBS[38]. The participants were first required to stand unsupported before a low stool (18 cm × 40 cm × 60 cm). Next, they were instructed to alternately put their entire right and left feet onto the stool as fast as possible. The time required to complete eight steps with each leg was recorded using a stopwatch. The AST was previously found to show good test-retest reliability (ICC = 0.78) and to discriminate recurrent fallers from non-recurrent fallers among community-dwelling older adults [21].

Sample size calculation

We estimated that a sample size of 30 would be required to achieve 90% power at a confidence level of .05 to detect strong test-retest reliability (ICC ≥ 0.90). Therefore, to account for a potential 20% drop-out rate, we aimed to recruit 38 participants.

Statistical analyses

A two-way mixed absolute agreement model of the ICC was used to analyze the test-retest reliability of the step tests, and a two-way random consistency model of the ICC was used to analyze the inter-rater reliability of the step tests [46]. Furthermore, a single-measures

model was applied to analyze the reliability of the CSRTT and AST, and an average-measures model was used to evaluate the reliability of the FSST and MSLT. ICCs of less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.9 were regarded as poor, moderate, good, and excellent reliability, respectively [46].

Pearson correlation coefficients (*r*) were used to analyze the correlations between the step test performances and physical and cognitive function and determine the concurrent validity of the step tests. Correlation coefficients between 0.3 and 0.6 and greater than 0.6 were regarded as moderate and strong correlations, respectively [47]. Independent *t*-tests were conducted to compare the step test performances of the participants in different subgroups (i.e., those who walked with a stick vs. those who walked without a stick, those who had a major neurocognitive impairment vs. those who did not) and determine the discriminative validity of the step tests.

The MDCs of a clinical test refers to the magnitude of change required to indicate an "actual" change in the individuals' performance beyond measurement error and individual variabilities [48]. The MDCs of the step tests were calculated as follows.

(1) The standard error of measurement (SEM) was calculated using the following formula [48]:

$$SEM = sd \times \sqrt{(1 - ICC)}$$

Where "sd" is the standard deviation of the step test performances, and "ICC" is the ICC indicating the test-retest reliability.

(2) The absolute MDC at a 95% confidence interval (MDC_{95}) was calculated using the following formula [48,49]:

$$MDC_{95} = SEM \times 1.96 \times \sqrt{2}$$

Where 1.96 represents the z score at the 95% confidence interval from a normal distribution, and the square root of 2 adjusts for repeated measurement errors.

(3) The relative MDC at a 95% confidence interval (MDC_{95}%) was calculated using the following formula [48]:

$MDC_{95}\% = (MDC_{95}/score) \times 100$

where "score" is the score on a step test.

All the analyses of concurrent and discriminative validity and MDCs were based on the participants' step test performances on the first testing occasion conducted by rater A. SPSS software (version 28.0) was used to perform the data analyses. A significance level of .05 was applied to all analyses.

Results

Characteristics

Figure 1 shows the participant recruitment, study procedures, and participants who have completed the study. Forty participants fulfilled our study criteria and were recruited. Thirty-seven participants completed the baseline assessments and the step tests. Three participants withdrew from the study because of the physical distancing measures implemented during the coronavirus 2019 pandemic or due to health problems unrelated to the assessments. Table 1 shows the characteristics and physical and cognitive measurements of the participants.

Test-retest and inter-rater reliability

The FSST, CSRTT, and MSLT showed good-to-excellent test–retest reliability (ICC = 0.83–0.96), whereas the AST exhibited modest test–retest reliability (ICC = 0.70). All step tests showed good-to-excellent inter-rater reliability (ICC = 0.75–0.95) (Table 2).

Concurrent validity

FSST performances were moderately to highly correlated with the scores of the HK-MoCA (r = -0.55), 2MWT (r = -0.42), 30STS (r = -0.50), and CDAD (r = -0.62) (Table 3).

CSRTT performances were moderately to strongly correlated with the scores of the HK-MoCA (r = -0.60), 2MWT (r = -0.54), 10mWT (r = -0.46), 30STS (r = -0.49), and CDAD (r = -0.67). MSLT performances were consistently correlated with the scores of the 2MWT (r = 0.53–0.70), 30STS (r = 0.38–0.60), and CDAD (r = 0.32–0.37). AST performances were moderately to highly correlated with the scores of the HK-MoCA (r = -0.60), 10mWT (r = -0.45), 30STS (r = -0.55), CDAD (r = -0.59), and 2MWT (r = -0.64).

Discriminative validity

The participants who used a stick had significantly shorter side step lengths in the MSLT (p = .014) and took a significantly longer time to complete the AST (p = .046) than those who did not use a stick to walk (Table 4). The participants who were regarded as having a major neurocognitive impairment took a significantly longer time to complete the FSST (p = .004), CSRTT (p = .005), and AST (p = .002) than those who did not.

Minimal detectable change

Supplementary Table 1 shows the SEM, MDC_{95}, and MDC_{95}% of the step tests. The $MDCs_{95}$% of the step tests ranged from 15% to 74%.

Discussion

This study is the first to investigate the psychometric properties of step tests in older adults with dementia. Our findings show that the FSST, CSRTT, and MSLT had good-to-excellent test-retest and inter-rater reliability. In contrast, the AST had fair test-retest reliability and good inter-rater reliability. In addition, the step tests had moderate-to-strong correlations with different physical and cognitive measures. The step tests could differentiate our sample with different ambulatory and cognitive statuses. These results suggest that the FSST, CSRTT, and MSLT can be used to evaluate the stepping performance of older adults with dementia. Nevertheless, the MDCs of the FSST, CSRTT, and AST were generally high,

suggesting that changes in the test results must be large to indicate actual changes in individuals' stepping performance in the real world.

The FSST, CSRTT, and MSLT had good-to-excellent test-retest and inter-rater reliability. These findings are similar to those reported in other populations, such as healthy older adults [18–20,32] and people with hip osteoarthritis [50], vestibular disorder [31], stroke [51–54], and Parkinson's disease [26]. However, the AST had only a modest test-retest and good-to-excellent inter-rater reliability. The lower confidence interval limit of the test-retest ICC was less than 0.5 (95% CI = 0.49–0.83), indicating that the true test-retest reliability may be poor [46]. The test-retest reliability of the AST in our sample was also less than that of previous studies on healthy older adults [55] and people with stroke[29]. In contrast, the inter-rater reliability of the AST in our sample is similar to that reported in a previous study on people with stroke [29]. Thus, the AST may not be a reliable tool for evaluating the stepping performance of older adults with dementia.

This study found moderate-to-strong correlations between the step test performances and general mobility, lower limb muscle strength, gait speed, and daily functioning. Furthermore, the FSST, CSRTT, and AST performances were moderately correlated with global cognition. Previous studies have shown moderate correlations between the step tests and measures of physical performance in healthy older adults [18,20] and people with Parkinson's disease [26,27], stroke [29,51–54], and multiple sclerosis [28]. However, the strong correlation between the FSST and the daily functioning measure in the current study is in contrast to another study, which showed no significant correlation between the FSST and a functional measure in people with vestibular disorder [31]. Moreover, the current study is the first to demonstrate a significant relationship between FSST, CSRTT, and AST performances and global cognition in older adults with dementia. One previous study showed that CSRTT performance was weakly associated with global cognition in healthy older adults (r = -0.29) [19]. The results indicate that the FSST, CSRTT, and AST require older adults with dementia to use both physical and

cognitive functions to perform. These step tests may evaluate cognitive function in addition to physical performance, particularly in older adults with dementia [56].

Older adults with dementia who walked without a walking stick performed significantly better in the MSLT and AST than those with a walking stick. This suggests that MSLT side step length and AST may be effective in distinguishing between older adults with dementia who have different ambulatory statuses. Similarly, it was previously suggested that the MSLT could be used to differentiate between people with and without chronic stroke[53] and an increased risk of falls [19,21,26,28,30]. Furthermore, the CSRTT and AST were also found in previous studies to be effective in predicting future falls among community-dwelling older adults [19,55]. Future studies using a prospective design and a larger sample are required to determine whether the step tests are equally effective in identifying fallers among older adults with dementia.

Compared with older adults with dementia who had less impaired global cognition, those with severely impaired global cognition took significantly more time to complete the FSST, CSRTT, and AST. This study is the first to show that these step tests can differentiate between people with different levels of cognitive impairment. The ability of step tests to differentiate older adults with different severity of cognitive impairment and deficits in selected cognitive domains (e.g., executive function) warrants further investigation.

The current findings of MDC_{95} provide useful information for clinicians to determine what magnitudes of changes in step tests are required to detect true changes in the stepping performance of older adults with dementia in real life. Furthermore, the MDC_{95}% results enable us to compare the required changes across different measures. We found that the MDC_{95}% values of the FSST, CSRTT, and AST were greater (39%–74%) than those of the MSLT (15% – 23%), as well as those of common walk tests [14,44] and other physical measures previously reported in the same population [36,57]. These findings suggest that the FSST, CSRTT, and AST may be less sensitive in identifying changes in the stepping performance of older adults with dementia. Older adults with advanced age and cognitive

impairment are expected to exhibit increased variability in their performances in physical measures [44,58]. We estimate that the variability of the step test performances in people with significant physical and cognitive deficits (e.g., older adults with dementia) may be further increased when these tests demand additional cognitive resources to complete. Our correlation analyses support this speculation, showing that these three-step tests may require older adults with dementia to use not just physical resources but also cognitive resources. Clinicians using the FSST, CSRTT, and AST to evaluate older adults with dementia must carefully examine the variability and sensitivity of these tests.

Although step tests and walk tests (e.g., Groningen Meander Walk Test and 6-meter Walk Test) share similar purposes and procedures – they both assess individuals' ability to take a series of steps and evaluate the mobility and motor function in older adults, step tests may provide additional values over walk tests in assessing older adults with dementia. First, the movements involved in some step tests (e.g., steps and lunges in different directions in the FSST and MSLT) are expected to challenge the lower limb strength and balance of older adults. These step tests may be more sensitive in identifying the related physical deficits in older adults than walk tests. Second, our findings show that the performance of older adults with dementia in step tests, particularly the FSST and CSRTT, is significantly associated with their cognitive function. Thus, step tests may be useful in evaluating not only discrete physical or cognitive function but also cognitive-motor performance in older adults with dementia. Nevertheless, further studies are required to determine whether step tests can predict health outcomes (e.g., falls, functional dependence, mortality) in this population.

Study Limitations

Our participants were relatively old and had comorbidities. Thus, our findings may not be generalizable to young older adults without comorbidities. We used the HK-MoCA score as one of the inclusion criteria. Some participants may have a cognitive impairment but not necessarily a diagnosis of dementia. We also did not recruit older adults with a specific type of dementia. The present findings provide an overview of the stepping performance of a

group of older adults with dementia and significant cognitive impairment. However, we are unable to determine whether the participants with different dementia subtypes performed differently in the step tests. A previous study showed that older adults with different dementia subtypes may have different walking performances and levels of fall risk [5]. To be able to generalize the data to type-, stage- and severity-specific types of dementia, future work using a large sample of older adults with different types and severity of dementia is required. The sample size was small and not sufficiently powered to allow comparisons of the step test performances of different subgroups and analyses of the correlations between the step tests and physical and cognitive measures. The small sample size and the large variance in the stepping performance of the participants may also affect the interpretation of the MDCs.

Conclusion

Our findings show that the FSST, CSRTT, and MSLT are reliable and valid tests for evaluating the stepping performance of older adults with dementia. These step tests achieved good to excellent test–retest and inter-rater reliability and had moderate-to-strong correlations with different physical and cognitive outcome measures. Moreover, the step tests could distinguish between older adults with dementia with different ambulatory statuses and severities of neurocognitive impairment. Nevertheless, clinicians need to be aware of the variability and sensitivity of these step tests when evaluating their stepping performances. The cognitive demand of the step tests and their ability to identify fallers in older adults with dementia need to be further investigated.

Acknowledgment

We want to thank the participants and their main caregivers in this project and the staff at multiple community centers for older adults who helped coordinate the recruitment and arranged the venue for the assessment sessions.

The materials of this manuscript have never been presented before in any circumstance. This feasibility trial was funded by The Hong Kong Polytechnic University (Project ID: P0030116).

Figure legends

- Figure 1. Participant recruitment and study procedure.
- Supplementary Figure 1. Set-up of (a) Four Square Step Test, (b) Choice Stepping Reaction Time Test, and (c) Maximum Step Length Test (dimensions in centimeters).

References

1. Fisher SR, Graham JE, Ottenbacher KJ, Deer R, Ostir GV. Inpatient Walking Activity to Predict Readmission in Older Adults. *Archives of Physical Medicine and Rehabilitation*. 2016;97(9, Supplement):S226-S231. doi:10.1016/j.apmr.2015.09.029

2. Verghese J, Holtzer R, Lipton RB, Wang C. Mobility stress test approach to predicting frailty, disability, and mortality in high-functioning older adults. *Journal of the American Geriatrics Society*. 2012;60(10):1901-1905. doi:10.1111/j.1532-5415.2012.04145.x

3. Kikkert LHJ, Vuillerme N, van Campen JP, Hortobágyi T, Lamoth CJ. Walking ability to predict future cognitive decline in old adults: A scoping review. *Ageing Research Reviews*. 2016;27:1-14. doi:10.1016/j.arr.2016.02.001

4. Kueper JK, Speechley M, Lingum NR, Montero-Odasso M. Motor function and incident dementia: a systematic review and meta-analysis. *Age and Ageing*. 2017;46(5):729-738. doi:10.1093/ageing/afx084

5. Allali G, Launay CP, Blumen HM, et al. Falls, Cognitive Impairment, and Gait Performance: Results From the GOOD Initiative. *Journal of the American Medical Directors Association*. 2017;18(4):335-340. doi:10.1016/j.jamda.2016.10.008

6. Taylor ME, Delbaere K, Mikolaizak AS, Lord SR, Close JCT. Gait parameter risk factors for falls under simple and dual task conditions in cognitively impaired older people. *Gait & Posture*. 2013;37(1):126-130. doi:10.1016/j.gaitpost.2012.06.024

7. Sterke CS, van Beeck EF, Looman CWN, Kressig RW, van der Cammen TJM. An electronic walkway can predict short-term fall risk in nursing home residents with dementia. *Gait & Posture*. 2012;36(1):95-101. doi:10.1016/j.gaitpost.2012.01.012

8. Dyer AH, Lawlor B, Kennelly SP. Gait speed, cognition and falls in people living with mild-to-moderate Alzheimer disease: data from NILVAD. *BMC Geriatrics*. 2020;20(1):117. doi:10.1186/s12877-020-01531-w

9. McGough EL, Logsdon RG, Kelly VE, Teri L. Functional mobility limitations and falls in assisted living residents with dementia : Physical performance assessment and quantitative gait analysis. *Journal of Geriatric Physical Therapy*. 2013;36(2):78-86. doi:10.1519/JPT.0b013e318268de7f

10. Toots A, Rosendahl E, Lundin-Olsson L, Nordström P, Gustafson Y, Littbrand H. Usual Gait Speed Independently Predicts Mortality in Very Old People: A Population-Based Study. *Journal of the American Medical Directors Association*. 2013;14(7):529.e1-529.e6. doi:10.1016/j.jamda.2013.04.006

11. Berg WP, Alessio HM, Mills EM, Tong C. Circumstances and consequences of falls in independent community-dwelling older adults. *Age and ageing*. 1997;26(4):261-268. doi:10.1093/ageing/26.4.261

12. Robinovitch SN, Feldman F, Yang Y, et al. Video capture of the circumstances of falls in elderly people residing in long-term care: an observational study. *The Lancet*. 2013;381(9860):47-54. doi:10.1016/S0140-6736(12)61263-X

13. Lord SR, Menz HB, Sherrington C. Home environment risk factors for falls in older people and the efficacy of home modifications. *Age and Ageing*. 2006;35(suppl_2):ii55-ii59. doi:10.1093/ageing/afl088

14. Chan WLS, Pin TW. Reliability, validity and minimal detectable change of 2-minute walk test, 6-minute walk test and 10-meter walk test in frail older adults with dementia. *Experimental Gerontology*. 2019;115:9-18. doi:10.1016/j.exger.2018.11.001

15. Brown LA, Shumway-Cook A, Woollacott MH. Attentional Demands and Postural Recovery: The Effects of Aging. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 1999;54(4):M165-M171. doi:10.1093/gerona/54.4.M165

16. Taylor ME, Lord SR, Delbaere K, Kurrle SE, Mikolaizak AS, Close JCT. Reaction Time and Postural Sway Modify the Effect of Executive Function on Risk of Falls in Older People with Mild to Moderate Cognitive Impairment. *American Journal of Geriatric Psychiatry*. 2017;25(4):397-406. doi:10.1016/j.jagp.2016.10.010

17. Taylor ME, Ketels MM, Delbaere K, Lord SR, Mikolaizak AS, Close JCT. Gait impairment and falls in cognitively impaired older adults: an explanatory model of sensorimotor and neuropsychological mediators. *Age and Ageing*. 2012;41(5):665-669. doi:10.1093/ageing/afs057

18. Dite W, Temple VA. A clinical test of stepping and change of direction to identify multiple falling older adults. *Archives of Physical Medicine and Rehabilitation*. 2002;83(11):1566-1571. doi:10.1053/apmr.2002.35469

19. Delbaere K, Gschwind Y, Sherrington C, Barraclough E, Garrués-Irisarri M, Lord S. Validity and reliability of a simple 'low-tech' test for measuring choice stepping reaction time in older people. *Clinical Rehabilitation*. 2016;30(11):1128-1135. doi:10.1177/0269215515613422

20. Medell JL, Alexander NB. A Clinical Measure of Maximal and Rapid Stepping in Older Women. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 2000;55(8):M429-M433. doi:10.1093/gerona/55.8.M429

21. Tiedemann A, Shimada H, Sherrington C, Murray S, Lord S. The comparative ability of eight functional mobility tests for predicting falls in community-dwelling older people. *Age and Ageing*. 2008;37(4):430-435. doi:10.1093/ageing/afn100

22. Rogers MW, Creath RA, Gray V, et al. Comparison of Lateral Perturbation-Induced Step Training and Hip Muscle Strengthening Exercise on Balance and Falls in Community-Dwelling Older Adults: A Blinded Randomized Controlled Trial. *The Journals of Gerontology: Series A*. 2021;76(9):e194-e202. doi:10.1093/gerona/glab017

23. Shigematsu R, Okura T, Nakagaichi M, et al. Square-Stepping Exercise and Fall Risk Factors in Older Adults: A Single-Blind, Randomized Controlled Trial. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 2008;63(1):76-82. doi:10.1093/gerona/63.1.76

24. Nnodim JO, Strasburg D, Nabozny M, et al. Dynamic Balance and Stepping Versus Tai Chi Training to Improve Balance and Stepping in At-Risk Older Adults. *Journal of the American Geriatrics Society*. 2006;54(12):1825-1831. doi:10.1111/j.1532-5415.2006.00971.x

25. Schoene D, Lord SR, Delbaere K, Severino C, Davies TA, Smith ST. A Randomized Controlled Pilot Study of Home-Based Step Training in Older People Using Videogame Technology. Berthold HK, ed. *PLoS ONE*. 2013;8(3):e57734. doi:10.1371/journal.pone.0057734

26. Duncan RP, Earhart GM. Four square step test performance in people with Parkinson disease. *Journal of neurologic physical therapy : JNPT*. 2013;37(1):2-8. doi:10.1097/NPT.0b013e31827f0d7a

27. Duncan RP, McNeely ME, Earhart GM. Maximum Step Length Test Performance in People With Parkinson Disease. *Journal of Neurologic Physical Therapy*. 2017;41(4):215-221. doi:10.1097/NPT.0000000000000201

28. Tijsma M, Vister E, Hoang P, Lord SR. A simple test of choice stepping reaction time for assessing fall risk in people with multiple sclerosis. *Disability and Rehabilitation*. 2017;39(6):601-607. doi:10.3109/09638288.2016.1148784

29. Chung MML, Chan RWY, Fung YK, et al. Reliability and validity of alternate step test times in subjects with chronic stroke. *Journal of Rehabilitation Medicine*. 2014;46(10):969-974. doi:10.2340/16501977-1877

30. Dite W, Connor HJ, Curtis HC. Clinical Identification of Multiple Fall Risk Early After Unilateral Transtibial Amputation. *Archives of Physical Medicine and Rehabilitation*. 2007;88(1):109-114. doi:10.1016/j.apmr.2006.10.015

31. Whitney SL, Marchetti GF, Morris LO, Sparto PJ. The Reliability and Validity of the Four Square Step Test for People With Balance Deficits Secondary to a Vestibular Disorder. *Archives of Physical Medicine and Rehabilitation*. 2007;88(1):99-104. doi:10.1016/j.apmr.2006.10.027

32. Goldberg A, Schepens S, Wallace M. Concurrent validity and reliability of the maximum step length test in older adults. *Journal of Geriatric Physical Therapy*. 2010;33(3):122-127. doi:10.1097/JPT.0b013e3181eda302

33. Suttanon P, Hill KD, Dodd KJ, Said CM. Retest reliability of balance and mobility measurements in people with mild to moderate Alzheimer's disease. *International Psychogeriatrics*. 2011;23(7):1152-1159. doi:10.1017/S1041610211000639

34. Wong A, Law LSN, Liu W, et al. Montreal cognitive assessment: One cutoff never fits all. *Stroke*. 2015;46(12):3547-3550. doi:10.1161/STROKEAHA.115.011226

35. World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. *Bull World Health Organ*. 2001;79(4):373-374.

36. Blankevoort CG, van Heuvelen MJG, Scherder EJA. Reliability of Six Physical Performance Tests in Older People With Dementia. *Physical Therapy*. 2013;93(1):69-78. doi:10.2522/ptj.20110164

37. Jones CJ, Rikli RE, Beam WC. A 30-s Chair-Stand Test as a Measure of Lower Body Strength in Community-Residing Older Adults. *Research Quarterly for Exercise and Sport*. 1999;70(2):113-119. doi:10.1080/02701367.1999.10608028

38. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. *Canadian journal of public health = Revue canadienne de sante publique*. 1992;83 Suppl 2:S7-11.

39. Muir-Hunter SW, Graham L, Odasso MM. Reliability of the berg balance scale as a clinical measure of balance in community-dwelling older adults with mild to moderate alzheimer disease: A pilot study. *Physiotherapy Canada*. 2015;67(3):255-262. doi:10.3138/ptc.2014-32

40. Mok CCM, Siu AMH, Chan WCC, Yeung KMM, Pan PCC, Li S wah W. Functional disabilities profile of Chinese elderly people with Alzheimer's disease - A validation study on the Chinese version of the disability assessment for dementia. *Dementia and Geriatric Cognitive Disorders*. 2005;20(2-3):112-119. doi:10.1159/000086612

41. Gélinas I, Gauthier L, McIntyre M, Gauthier S. Development of a Functional Measure for Persons With Alzheimer's Disease: The Disability Assessment for Dementia. *The American Journal of Occupational Therapy*. 1999;53(5):471-481. doi:10.5014/ajot.53.5.471

42. Chan WLS, Pin TW. Practice effect and cueing of 2-minute walk test, 6-minute walk test and 10-meter walk test in frail older adults with and without dementia - Recommendations to walk tests protocols. *Experimental Gerontology*. 2019;124:110648. doi:10.1016/j.exger.2019.110648

43. Trautwein S, Barisch-Fritz B, Scharpf A, et al. Recommendations for assessing motor performance in individuals with dementia: suggestions of an expert panel – a qualitative

approach. *European Review of Aging and Physical Activity*. 2019;16(1):5. doi:10.1186/s11556-019-0212-7

44. Chan WLS, Cheung YT, Lee YW, Teo AM, Wo HK, Wong Y. Reliability, Validity, and Minimal Detectable Change of the Backward Walk Test in Older Adults With Dementia. *Journal of Geriatric Physical Therapy*. 2022;45(3):145-153. doi:10.1519/JPT.0000000000000306

45. Lord SR, Fitzpatrick RC. Choice Stepping Reaction Time: A Composite Measure of Falls Risk in Older People. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 2001;56(10):M627-M632. doi:10.1093/gerona/56.10.M627

46. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*. 2016;15(2):155-163. doi:10.1016/j.jcm.2016.02.012

47. Andresen EM. Criteria for assessing the tools of disability outcomes research. *Archives of Physical Medicine and Rehabilitation*. 2000;81(12 SUPPL. 2):15-20. doi:10.1053/apmr.2000.20619

48. Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek ALM. Smallest real difference, a link between reproducibility and responsiveness. *Quality of Life Research*. 2001;10(7):571-578. doi:10.1023/A:1013138911638

49. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research*. 2005;19(1):231-240. doi:10.1519/15184.1

50. Choi YM, Dobson F, Martin J, Bennell KL, Hinman RS. Interrater and Intrarater Reliability of Common Clinical Standing Balance Tests for People With Hip Osteoarthritis. *Physical Therapy*. 2014;94(5):696-704. doi:10.2522/ptj.20130266

51. Blennerhassett JM, Jayalath VM. The Four Square Step Test is a Feasible and Valid Clinical Test of Dynamic Standing Balance for Use in Ambulant People Poststroke. *Archives of Physical Medicine and Rehabilitation*. 2008;89(11):2156-2161. doi:10.1016/j.apmr.2008.05.012

52. Goh EY, Chua SY, Hong SJ, Ng SS. Reliability and concurrent validity of four square step test scores in subjects with chronic stroke: A pilot study. *Archives of Physical Medicine and Rehabilitation*. 2013;94(7):1306-1311. doi:10.1016/j.apmr.2013.01.027

53. Ng SSM, Tse MMY, Kwong PWH, et al. Reliability of the Maximal Step Length Test and Its Correlation with Motor Function in Chronic Stroke Survivors. *BioMed Research International*. 2018;2018:1-8. doi:10.1155/2018/6985963

54. Pardo V, Knuth D, McDermott B, Powell J, Goldberg A. Validity, reliability and minimum detectable change of the maximum step length test in people with stroke. *Journal of the Neurological Sciences*. 2013;325(1):74-78. doi:10.1016/j.jns.2012.11.021

55. Tiedemann A, Lord SR, Sherrington C. The Development and Validation of a Brief Performance-Based Fall Risk Assessment Tool for Use in Primary Care. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 2010;65A(8):896-903. doi:10.1093/gerona/glq067

56. Pelicioni PHS, Lord SR, Sturnieks DL, Halmy B, Menant JC. Cognitive and Motor Cortical Activity During Cognitively Demanding Stepping Tasks in Older People at Low and High Risk of Falling. *Frontiers in Medicine*. 2021;8:1016. doi:10.3389/fmed.2021.554231

57. Braun T, Thiel C, Schulz RJ, Grüneberg C. Reliability of mobility measures in older medical patients with cognitive impairment. *BMC Geriatrics*. 2019;19(1):20. doi:10.1186/s12877-019-1036-z

58. Fox B, Henwood T, Keogh J, Neville C. Psychometric viability of measures of functional performance commonly used for people with dementia. *JBI Database of Systematic Reviews and Implementation Reports*. 2016;14(8):115-171. doi:10.11124/JBISRIR-2016-003064