7-3-2020

Reliable and Valid Measures for the Clinical Assessment of Balance and Gait in Older Adults with Dementia: A Systematic Review

Jacqueline Baker
Western University, jbake6@uwo.ca

Danielle de Laat
Western University, ddelaat2@uwo.ca

Emily Kruger
Western University, ekruger6@uwo.ca

Sarah McRae
Western University, smcrae9@uwo.ca

Sabrina Trung
Western University, strung@uwo.ca

See next page for additional authors

Reliable and valid measures for the clinical assessment of balance and gait in older adults with dementia: A systematic review

Jacqueline Baker,[1] Danielle de Laat,[1] Emily Kruger,[1] Sarah McRae,[1] Sabrina Trung,[1] Carly Zottola,[1] Humberto Omaña,[2] Susan W Hunter[1, 2]

1. School of Physical Therapy, University of Western Ontario, London, Ontario;
2. Faculty of Health Sciences, University of Western Ontario, London, Ontario.

Corresponding Author:

Dr. Susan W. Hunter

University of Western Ontario

School of Physical Therapy

Elborn College

London, ON, Canada, N6G 1H1

Phone: 519-661-2111 ext 88845

Email: susan.hunter@uwo.ca

Reliable and valid measures for the clinical assessment of balance and gait in older adults with dementia: A systematic review

ABSTRACT

Purpose: To systematically review the literature and determine clinical balance or gait tests appropriate for use in older adults with dementia.

Materials and Methods: Databases PubMed, EMBASE and CINAHL were searched (inception-April 2020). Inclusion criteria: participants were diagnosed with dementia and reliability or validity was evaluated for clinical balance or gait tests. Exclusion criteria: not published in English or unable to extract data.

Results: Twenty-two studies evaluating 35 clinical balance or gait tests were included. For community-dwelling individuals, the Modified Clinical Test of Sensory Interaction in Balance, gait velocity, Step Test and the Timed Up and Go had good relative and absolute reliability. For institutional-dwelling individuals the Berg Balance Scale, the Groningen Meander Walk Test, and 6-Meter Walk Test had excellent relative and absolute reliability. The Groningen Meander Walk Test was the only recommended test that has been validated.

Conclusions: Reliable clinical tests of balance and gait for use in people with dementia exist, yet few have been validated. Additionally, to comprehensively assess balance, clinicians must utilize multiple tests. Future research should evaluate the psychometric properties of clinical balance and gait tests so as to identify those appropriate for use in people with dementia across setting, severity and dementia type.

KEYWORDS: dementia, postural balance, gait, reproducibility of results, systematic review

INTRODUCTION

Dementia is associated with overall cognitive decline, decreased physical function and a loss of mobility that increases the risk for falls.[1] People with dementia have an annual falls risk of 60-80%, twice that of the cognitively normal older adult,[2] and have a higher risk of major fall-related injuries, such as hip fractures.[3] Balance and gait impairment are major risk factors for falls in older adults with dementia living in the community and in an institutional setting.[4] There are approximately 43.8 million people in the world currently living with dementia and this number is expected to rise to 100 million by the year 2050.[5] Physical therapists rely on standardized clinical tests of balance and gait to identify deficits and develop interventions aimed at reducing mobility impairments, falls and improving quality of life in this population. Therefore, an understanding of which clinical balance and gait tests are appropriate to use in people living with dementia is warranted.

Numerous clinical tests of balance and gait currently exist; most which have been developed in cognitively healthy older adults.[6,7] However, these tests may not be appropriate for use in people with dementia due to disease-specific cognitive deficits. Specifically, an ability to understand complex instructions and memory to execute multi-step commands are common requirements for clinical tests of balance and gait. As a consequence of limited cognitive abilities, people with dementia often demonstrate reduced balance and gait performance that puts into question how reliable and valid these clinical tests are. More importantly, relying on clinical tests with no known or poor psychometric properties affects healthcare professionals' ability to accurately quantify ability, track progress and initiate intervention.[8]

The previous systematic review by Bossers et al.[9] evaluated randomized controlled trials to provide recommendations for neuropsychological and physical function tests for use in people

with dementia. Similarly and more recently, authors Trautwein et al.[10] provided recommendations of motor assessments that have previously been used in randomized controlled trials of physical activity interventions in people with dementia. Importantly, these recommendations are seven and close to two years old, respectively. This area of work merits updating to include a broader depiction of work published since these systematic reviews to further inform clinical test selection recommendations. Likewise, Van Ooteghem et al.[11] examined the feasibility of standardized clinical tests assessing mobility in those with advanced dementia. As this study focused on advanced dementia, the state of the literature on clinical tests of balance and gait suitable to use in people with mild to moderate disease was not reviewed recently.

To the authors' knowledge, no systematic review to date has examined the psychometric properties of clinical tests of balance and gait in older adults with dementia across healthcare settings. Thus, the primary aim of this study was to systematically review the literature to determine which standardized clinical tests of balance and gait are appropriate for use in older adults with dementia, based on reported psychometric properties for reliability and validity.

MATERIALS AND METHODS

Data Sources and Searches

Articles published in databases PubMed, EMBASE, and CINAHL were all searched from their inception to April 2020. Articles were pulled from each database in duplicate by trained reviewers working in pairs (JB, DdL, EK, SM, ST, CZ) and within the same week of each other. The literature searches were developed in conjunction with a research librarian. The following keywords were used in combination and without any database filtering or restrictions: dementia

(OR delirium OR amnestic OR cognitive disorders), balance (OR postural OR gait), psychometric values (predictive value of tests OR reproducibility of results OR test-retest reliability OR psychometrics OR instruments validation OR validity OR reliability OR pilot studies). (A sample search strategy is included in Supplementary Table 1). Moreover, reference lists of the extracted articles and existing reviews related to the topic of interest were hand searched for additional studies not captured in the electronic searches. The systematic review was designed to adhere to The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines,[12] and was registered in The International Prospective Register of Systematic Reviews (PROSPERO; #CRD42020146944).

Study Selection

Studies were included if they met the following criteria: 1) all the participants had a diagnosis of dementia, 2) the study used standardized clinical tests for balance or gait, and 3) the study reported psychometric values related to reliability (i.e., inter-rater reliability, intra-rater reliability, test-retest reliability, standard error of measurement, minimal detectable change) or validity (e.g., content, concurrent, predictive). Studies were excluded if: 1) they were not published in English, 2) unable to extract data, 3) the study was a systematic or scoping review or meta-analysis, or 4) did not assess reliability or validity.

After the literature search, all study duplicates were removed. Working in pairs, six of the authors (JB, DdL, EK, SM, ST, CZ), independently reviewed the abstracts for inclusion criteria. Upon meeting initial inclusion criteria, a full-text examination of each article was completed. Consensus between reviewers was a requirement for articles to be included in the final analysis and data extraction. A third reviewer (SWH) was introduced to resolve any disagreements.

Data Extraction and Methodological Reporting Quality Assessment

The following data were extracted in duplicate by authors in pairs (JB, DdL, EK, SM, ST, CZ) from the studies included in the systematic review: authors, country, date of publication, study design, sample size, gender of participants, setting of recruitment, dementia diagnosis, diagnostic criteria for dementia diagnosis, balance and gait tools assessed and psychometric properties of reliability (i.e., inter-rater reliability, intra-rater reliability, test-retest, standard error of measurement (SEM), and minimal detectable change (MDC)) or validity (i.e., content, concurrent, predictive). Data extraction was completed using a standardized data extraction sheet piloted to meet the systematic review objectives. Study designs were reported as test-retest, cross-sectional or prospective in nature. Studies where the test was administered more than once over a period of time in which no change in the measure was expected to occur were considered test-retest. A cross-sectional study collected all data at one point in time.

Quality of reporting was evaluated using the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies (QOCCSS)[13] or the Quality Appraisal Tool for Studies of Diagnostic Reliability (QAREL)[14,15]. The QOCCSS is a 14-item appraisal tool focused on key concepts for evaluating the internal validity of observational studies, such as: selection bias, timeframe suitability, validity of exposure testing, assessor blinding, follow-up rate, and appropriate statistical analysis. The QAREL is an 11-item appraisal tool that covers seven key domains including: the spectrum of subjects and examiners, examiner blinding, the order effects of examination, the suitability of the time-interval between repeated measurements, appropriate test application and interpretation, and appropriate statistical analysis. Reliability studies were assessed using the QAREL, while studies containing measures of validity were assessed using the QOCCSS. Articles were split-up between pairs of reviewers and independently evaluated

(JB, DdL, EK, SM, ST, CZ). Consensus was required and any study quality scoring disagreement was resolved by a third reviewer (SWH).

Data Synthesis and Analysis

A clinically relevant depiction of the reliability and validity of standardized clinical tests of balance and gait used in people with dementia across healthcare settings was performed in the studies that met inclusion criteria. For each clinical test of balance identified in the review, the consensus paper by Sibley et al.[16] was used to identify the components of balance from the Systems Framework for Postural Control by Horak[17] that each test evaluated. There are nine components of postural stability named in the framework: anticipatory postural control, cognitive processing, dynamic stability, functional stability limits, motor system, reactive postural control, sensory strategies, static stability and verticality/orientation in space.

The following intraclass correlation coefficient (ICC) thresholds were used to interpret inter-rater, intra-rater and test-retest relative reliability values: an ICC >0.90 was deemed excellent, 0.80-0.89 was good, 0.70-0.79 was considered fair and an ICC <0.70 had questionable clinical value.[18] Although no absolute reliability thresholds currently exist to help guide interpretation of SEM or MDC, a smaller value indicates greater absolute reliability.[19] To interpret Pearson's (r) and Spearman's (ρ) correlation coefficients related to the assessment of concurrent validity, the following thresholds were used: ≥0.50 was deemed strong, 0.31-0.49 was moderate to strong, 0.11-0.30 was weak to moderate, and ≤0.10 was considered a non-existent relationship.[19]

RESULTS

There were 1325 unique articles identified and 58 full-text articles were screened for inclusion. (see Supplementary Table 2 for a summary of full text papers reviewed and excluded) Overall, 22 studies met the criteria for inclusion in this review.[8,20–40] (Figure 1) Sample sizes of studies ranged from 12-632 people. Nine studies[8,20,24,28,29,34–37] had samples of people with only Alzheimer's disease and the remaining articles had mixed samples and did specify the complement of dementia subtypes. Ten (45.5%) studies reported that criteria for dementia diagnosis was based on established guidelines.[8,20,25,26,28,29,34,35,37,38] Six[8,20,29,34,35,37] reported to have used the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) guidelines, three[25,28,38] the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR), and one[26] the Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Severity of dementia was reported in 14 (63.6%) studies.[8,20,22,24–26,28,29,34–39] Eight stated that their participants had mild to moderate dementia,[8,22,25,29,34–37] four reported moderate to severe dementia,[20,28,38,39] and two reported mild to severe dementia[24,26]. To assess for severity of dementia, eight studies[8,25,29,34–37,39] relied on the Mini-Mental State Examination (MMSE), Two[28,38] on the Global Deterioration Scale (GDS), one[22] on the Clinical Dementia Rating Scale (CDR), one[24] on the Functional Assessment Staging Scale (FAST), and two[20,26] did not specify how severity was assessed once the participants were recruited. Articles covered three settings: seven were community-dwelling,[8,29,31,34–37] nine were institution-dwelling[20–22,28,30,32,38–40] and six were a mixed-dwelling (sample was a mix of community and institution-dwelling older adults)[23–27,33].

Clinical Tests of Balance and Gait

There were 35 clinical tests of balance and gait evaluated in the reviewed papers. (Table 2) The following clinical tests for balance were evaluated across the papers: Balance Outcome

Measure for Elder Rehabilitation (BOOMER),[21] Berg Balance Scale (BBS – English,[8] Norwegian[22] and Korean[23]), Cooperative Studies of Intervention Techniques-4 Balance Test (FISCIT-4),[25] Functional Reach Test,[8,36] Limits of Stability Test,[36] Modified Clinical Test of Sensory Interaction in Balance (mCTSIB),[36] One Leg Balance Test,[29] Step Test,[36] Timed Up and Go Test (TUG – English[8,24,25,30,31] and Korean[23]), Tinetti Performance Oriented Mobility Assessment (POMA)[38]. There were dual-task test variants of the Timed Up and Go Test evaluated as well – TUG with motor task of carrying a glass of water (TUGmotor)[36] and TUG while counting backwards by threes (TUGcognitive)[36].

Gait was evaluated using the clinical tests of the Groningen Meander Walking Test (GMWT, English[40] and Korean[23,32]), Figure-of-Eight Walk Test (F8W),[25] 2-Minute Walking Test (2MWT),[30,33] and 6-Minute Walk Test (6MWT)[20,24,33]. Gait speed was evaluated using a stopwatch over different distances or 25-Foot Walk Test at usual pace,[20] 4-Meter Walk Test (4mWT),[23,30] 6-Meter Walk Test (6mWT), and 10-Meter Walking Test (10mWT),[33] at usual[22,25,31,] and fast pace[31]. Instrumented gait analysis collected a range of spatiotemporal parameters for single-task gait at usual pace: velocity,[24,28,35–37,39] cadence,[35,37,39] step length,[35,36] step width,[36] stride length,[28,35,37,39] stride width,[37] stride width variability,[37] support base,[35] toe in/out,[35], swing time,[35,39] stance time,[35] heel-to-heel base of support variability,[28] double support time variability;[28] and dual-task gait with secondary cognitive tasks: velocity[26].

Finally there were two clinical tests with balance and/or gait tasks that were included in a composite score: Physiological Profile Assessment (PPA)[34] and Short Physical Performance Battery (SPPB – English[21] and Norwegian[27]).

Reliability: Community-dwelling setting

Six studies evaluated reliability in the community-dwelling setting.[8,31,34–37] Twelve clinical tests of balance or gait were evaluated in this population. (Supplementary Table 3)

Inter-rater reliability was examined for the BBS, Functional Reach Test, and TUG.[8] All were found to have at least a fair inter-rater reliability with ICC values ranging from 0.72-0.98. Regarding test-retest reliability, the BBS,[8] Functional Reach Test,[8,36] mCTSIB,[36] single-task spatiotemporal gait parameters (velocity, cadence, step length, stride length, stride width, support base, toe in/out, swing time, stance time),[26,35,37] Step Test,[36] TUG,[8,31,36] and 6mWT[31] had at least good test-retest ICC values (0.80-0.97).

The SEM and MDC were calculated for 9 of the 12 clinical tests of balance and gait: BBS,[8] Functional Reach Test,[8,36] Limits of Stability,[36] mCTSIB,[36] single-task spatiotemporal gait parameters (velocity, cadence, stride length, stride width),[37] Step Test,[36] TUG,[8,36] TUGcognitive,[36] TUGmotor[36].

Overall, spatiotemporal gait parameters,[35,37] TUG,[8,36] Step Test[36] and the mCTSIB[36] were considered the best clinical tests in terms of having both good relative and absolute reliability. The BBS[8] and Functional Reach Test[8,36] had good relative reliability, but variable or poor absolute reliability. The PPA,[34] Limits of Stability Test[36] and Quick Turn Test[36] demonstrated both poor relative and absolute reliability, suggesting these clinical tests are unreliable in this patient population.

Reliability: Institutional-dwelling setting

Seven studies evaluated the reliability of twelve clinical tests of balance or gait in the institution-dwelling setting.[20–22,30,38–40] (Supplementary Table 4).

Inter-rater reliability was evaluated for five clinical tests of balance or gait in this setting. The BBS,[22] POMA,[38] 6mWT,[22] and 6MWT[20] were found to have excellent reliability (ICC >

0.90). The 25-Foot Walk Test had good-excellent inter-rater reliability (ICC range: 0.83-0.97).[20] Intra-rater reliability was investigated for two clinical tests of gait in one study;[20] the 6MWT[20] distance and gait velocity had at least fair-good reliability (ICC range: 0.75-0.90). Test-retest reliability was investigated for seven clinical tests of balance and gait. The GMWT (overall component),[40] spatiotemporal gait parameters (cadence, gait speed, stride velocity, swing time)[39], and the TUG[30] (cued) were all found to have excellent reliability.

The SEM and MDC was calculated for three clinical tests of balance and gait: BBS,[22] GMWT,[40] and the 6mWT[22].

The BBS,[22] GMWT,[40] and 6mWT[22] were the best clinical tests of balance and gait in terms of having excellent relative and absolute reliability. The POMA,[38]spatiotemporal gait parameters,[39] and the TUG[30] (cued) demonstrated excellent relative reliability but had no data regarding absolute reliability in this setting. Similarly, the 2MWT[30] (cued), and the 4MWT[30] (cued) had good relative reliability but no absolute reliability information. The 25-Foot Walk Test,[20] 6MWT,[20] BOOMER,[21] and the SPPB[21] had variable or poor relative and absolute reliability, making the reliability of these clinical tests for use in a clinical setting difficult to interpret.

Reliability: Mixed-dwelling setting

Six studies evaluated reliability for 15 clinical tests of balance or gait in the mixed-dwelling setting.[23–27,33] (Supplementary Table 5).

The Korean versions of BBS, GMWT, TUG, and 4mWT were all found to have at least good inter-rater reliability with ICC values ranging from 0.82 to 0.99.[23] The Inter-rater reliability was excellent for the cued 2MWT and 6MWT tests.[33] Regarding test-retest reliability, the 4-meter walk subtests of the SPPB Norwegian version,[27] the BBS Korean version,[23] dual-task gait

test with cognitive task,[26] F8W,[25] FISCIT-4,[25] GMWT Korean version,[23] TUG,[25] TUG Korean version,[23] 2MWT[33] (cued), 6MWT[33] (cued), 6mWT[25] , and 10mWT[33] (cued) all had at least a fair test-retest ICC value (ICC range: 0.75-0.99).

The SEM and MDC was calculated for 14 of the clinical tests of balance and gait: BBS Korean version,[23] F8W,[25] FISCIT-4,[25] gait velocity,[24] GMWT Korean version,[23] SPPB Norwegian version,[27] TUG,[24,25] TUG Korean version,[23] 4mWT Korean version,[23] 6mWT,[25] 2MWT,[33] 6MWT[24] (feet), 6MWT[33] (m), and the 10mWT[33].

Overall, the BBS Korean version,[23] GMWT Korean version,[23] TUG,[24,25] TUG Korean version[23], 2MWT[33] (cued), the 6MWT[33] (cued), and the 10mWT[33] (cued) were the best clinical tests in terms of having both excellent relative and absolute reliability. The SPPB (Norwegian version),[27] 4mWT Korean version[23] and 6mWT Korean version[23] have both good relative and absolute reliability. The F8W[25] had excellent relative reliability, but variable scores for absolute reliability. The FISCIT-4[25] had fair relative reliability and good absolute reliability. Gait velocity,[24] dual-task gait test with cognitive task[26] and 6MWT[24] (feet) had limited data demonstrating variable scores, making it difficult to interpret their reliability.

Validity: Across Settings

A total of seven studies[26,28,29,32,33,38,39] assessed the concurrent and predictive validity of eight clinical tests of balance and gait. (Supplementary Table 6).

In the community-dwelling setting, one study examined the predictive validity of the One-Leg Balance Test (OLB).[29] During a two year follow-up, individuals with abnormal OLB tests (less than 5 seconds) had significantly higher age- and sex-adjusted: 1) incidence of functional decline (loss of 0.5 points on the Katz's ADL scale), 2) incidence of "loss of walking

abilities" (loss of 0.5 points on the "walking ability" item of the Katz's ADL scale), 3) incidence of nursing home admission, and 4) risk of death per year than individuals with normal OLB tests.

Four studies in institutional-dwelling setting,[26,28,32,39] assessed the concurrent and predictive validity of three clinical tests of balance and gait. The Groningen Meander Walking Test had a significant and strong correlation to the Timed Up & Go (ρ=0.69).[32] Additionally, the spatiotemporal gait parameters of cadence, cycle double support, gait speed, stride length, stride length variability, stride velocity, and swing time variability were all found to have a significant, strong correlation to the Short Physical Performance Battery and modified Berg Balance Scale.[39]

Two studies assessed predictive validity in the institutional-dwelling setting.[28,38] Recommended cut-off points for the prediction of a fall within a 3-month period were calculated for the Tinetti Performance Oriented Mobility Assessment,[38] and the spatiotemporal gait parameters of: double support time variability, heel-to-heel base of support variability, stride length and velocity.[28] Importantly, the Tinetti Performance Oriented Mobility Assessment (HR: 1.08)[38] and each of the four gait parameters (OR: 1.19-1.53)[28] were significantly associated with the occurrence of a fall within a 3-month period in models adjusted for age, gender, and falls history.

In the mixed-dwelling setting, two studies examined the concurrent validity of four clinical tests of balance and gait.[26,33] Dual-task testing walking speed was reported to have a significant and strong correlation with the Timed Up & Go (r: -0.84 to -0.74), the Tinetti Performance Oriented Mobility Assessment (r: 0.62-0.80), and a weak to moderate strength correlation with the Mini-Mental State Examination (r: 0.28-0.37).[26] The cued 2MWT, 6MWT and 10mWT, had statistically significant weak to strong correlations to the Elderly Mobility

Scale (ρ: 0.27-0.43), the Berg Balance Scale (ρ: 0.35-0.49), and the Modified Barthel Index (ρ: 0.45-0.54).[33]

Components of Postural Control

Twelve of the clinical tests of balance evaluated in this review had consensus ratings. (Table 3) The following were the most common postural components assessed: motor system (12/12), anticipatory postural control (11/12), and dynamic stability (8/12). The components that were least evaluated were: verticality/orientation in space (0/12), reactive postural control (2/12), cognitive processing (2/12) and functional stability limits (3/12). All clinical tests of balance assessed a minimum of 3 postural control components, the most comprehensive clinical test of balance in the reviewed articles were the BBS (6/9) and the POMA (6/9).

Methodological Reporting Quality Assessment

In the community setting, the QAREL scores for six studies[8,31,34–37] ranged from 3-7/11 and the QOCCSS score for one study[29] was 11/14. In the institutional setting, the QAREL scores for seven studies[20–22,30,38–40] ranged from 4-7/11 and the QOCCSS score for four studies ranged from 5-10/14[28,32,38,39]. In the mixed setting, six studies [23–27,33] were assessed by the QAREL and scores ranged from 4-7/11, while one study[33] was assessed by the QOCCSS and had a score of 8/14. For the QOCCSS in the community and institutional settings, the included papers lacked information on whether the outcome assessors were blinded to the exposure status of participants. For the QAREL, more than half of the papers in the community and institutional settings did not report if the assessors had been blinded (5 items). Overall, there was no observable trend in the quality of papers between the community, institutional or mixed-dwelling settings, but scores indicated low to moderate methodological reporting quality.

DISCUSSION

The aim of this systematic review was to determine the standardized clinical tests of balance and gait that have appropriate psychometric properties of reliability and validity for use in people with dementia. For community-dwelling older adults with dementia, the clinical tests that demonstrated the most suitability with good relative and absolute reliability were spatiotemporal gait parameters (i.e., velocity, cadence, step length, stride length, stride width, support base (cm), swing time, toe in/out (degrees)), TUG, Step Test and the mCTSIB. For institutional-dwelling older adults with dementia, the BBS, GMWT and 6mWT demonstrated excellent absolute and relative reliability. Finally, in a mixed-dwelling population excellent absolute and relative reliability was established for the Korean versions of the BBS and GMWT, TUG (English and Korean version), and the cued 2MWT, 6MWT and 10mWT. Importantly, only the GMWT for the institutional-dwelling and the 2MWT, 6MWT and 10mWT in the mixed-dwelling settings were reported to have demonstrated validity in people living with dementia. (Table 4)

Bossers et al.'s systematic review evaluated literature published up to 2011 for the utility of neuropsychological and physical outcome measures for use in clinical trials among people with dementia.[9] The authors evaluated the reliability and validity of 10 tests of mobility and recommended the use of the POMA, 6mWT, TUG, and 6MWT. More recently, and up to 2018, Trautwein et al. examined the psychometric properties of motor assessments previously used in randomized controlled trials of physical activity in people with dementia.[10] The authors recommended the use of the Functional Reach Test, GMWT, BBS, POMA, TUG, spatiotemporal gait parameters, the Sit-to-Stand Test, and the 6MWT. However, and due to the objective of their systematic review, the psychometric properties of only 28 motor assessments were examined, 15

which encompassed balance or gait. Our review was not as restricted and as a result was able to identify 35 clinical tests of balance and gait, many published since 2011 that highlight this as an emerging area of research focus. Moreover, the majority of papers in our review included people with mild to moderate dementia, which fills a gap and complements the systematic review by Van Ooteghem et al. that focused on severe dementia.[11] Our review has demonstrated that the dwelling setting has an influence on the tests that can be recommended for use in clinical practice. For example, use of the BBS (Korean and English versions) was recommended for an institution setting, but the English version was not recommended for community-dwelling due to large absolute reliability values. The absence of a recommendation can also reflect a lack of evaluation of a tool across all settings. Therefore, further research is required to fully address the applicability of clinical test of balance and gait in all settings that physical therapists deliver rehabilitation to people living with dementia.

Clinical tests of balance and gait deemed appropriate in other patient populations should not be expected to be suitable for use in people with dementia. Research needs to invest in the development of new, or the modification of established, clinical tests that standardize and mitigate the challenges faced by clinicians when assessing balance and gait in this population. However, it is also important to note that the psychometric properties of any one given clinical test may be influenced by factors related to study design and execution. Thus, researchers are highly encouraged to follow strict recruitment efforts and reporting of participant characteristics aimed at reducing patient sample heterogeneity. Moreover, few of the clinical tests recommended in this review have been validated against other established clinical tests for use in people with dementia. Future research should consider the use of instrumented technology, such as force plate recordings of sway for balance and accelerometer recordings of spatiotemporal gait

parameters to help facilitate the validation process of clinical tests of balance and gait. Currently, there are no falls prevention guidelines that have been specifically developed for people with dementia and those for cognitively healthy older adults may not be fully relevant.[41] Therefore, another avenue for future research is to evaluate the predictive validity of test scores to future falls risk, as some of the standardized tests have required modification to be effective in people with dementia and diagnostic test properties from cognitively healthy older adults may not be valid.

No standardized clinical test of balance evaluated all nine components of postural stability as described by Horak (2006). A deficit in any one or a combination of these components will result in a different context-specific instability, leading to an increased risk of falls.[17] Clinical tests aid in the identification of deficits that can be addressed with rehabilitation and there is emerging evidence that exercise interventions can reduce falls risk and falls in community-dwelling adults with dementia.[42] The components of balance least examined were verticality, cognitive processing, functional stability limits, and reactive postural control. To comprehensively assess balance, clinicians have to be deliberate and utilize multiple different clinical tests of balance.[16] Future research is needed to develop and evaluate clinical tests of balance that focus on the components that are not addressed with recommended clinical tests in people with dementia.

Our systematic review identified many gaps in the literature for standardized clinical tests of balance and gait appropriate for use in people living with dementia. For example, few clinical tests were assessed in all settings or reported on the same psychometric properties. The majority of the studies reviewed focused their research on the reliability of clinical tests of balance and gait in people with mild to moderate dementia or did not specify severity. It is therefore

important to recognize that the recommendations of our review were based on reliability parameters as few validity studies exist. Meaning, our recommendations are a reflection of the current state of this field and may not translate to use in the severe stage of the disease regardless of living situation. Additionally, the authors would like to acknowledge that some of the recommended clinical tests of balance and gait require equipment and space that may not be available to all clinicians. As a result, healthcare professionals may need to accommodate using readily available tools (i.e., using a stopwatch and known distance to calculate gait speed), or supplement testing gaps using other objective (e.g., strength tests) or subjective (e.g., self-reported mobility) reports that were not part of the present systematic review.

This systematic review has several strengths worth highlighting. Our search was comprehensive in depth and breadth as three electronic databases were searched from their inception using search terms that covered the spectrum of dementia. Further, a methodological reporting quality review was performed to provide support for the strength of the evidence. There are several limitations of this review, however, that should be noted. Firstly, the review of the methodological reporting quality of the studies revealed that no study scored the best possible rating and most may be considered of low to moderate methodological reporting quality. Studies in this review did not consistently specify if assessors were blinded which can introduce bias. However, and although a number of the included studies demonstrated low methodological reporting quality, all studies included in this review used appropriate study designs, statistical analysis and interpretation of results; thus, all were considered in our recommendations. Interpretation of the results should also consider that many standardized clinical tests of balance and gait were only evaluated in one study; very few measures were evaluated across multiple studies or presented complete assessments of reliability and validity. In addition, we included

articles across the entire spectrum of dementia, but the classification of dementia severity and type was not consistent across studies and analysis of the results according to severity was limited in this review.

This systematic review has identified standardized clinical tests of balance and gait appropriate for community-dwelling, institutionalized, and mixed-dwelling settings for older adults with dementia. Clinical tests that assess balance and gait in healthy adults may not be reliable in those with dementia due to an increased difficulty to comprehend instructions or to repeat a sequence of tasks. As a consequence, the results may reflect poor cognitive function rather than physical ability. Therefore, changes in cognitive function that accompany the disease processes of dementia warrant the use of clinical tests that have established reliability and validity within samples of people living with dementia. The information from this review is important for clinicians as it ensures best-evidence informs the evaluation of balance and gait function that can be used to facilitate provision of interventions. More research is needed to evaluate the reliability and validity of clinical tests of balance and gait in people with dementia across care settings, severity and dementia type.

Funding details

None to disclose.

Disclosure of interest

The authors report no conflict of interest.

REFERENCES

1. Alzheimer Society of Canada, Canada AS of. *Rising Tide: The Impact of Dementia on Canadian Society*.; 2010.

2. Tinetti M, Speechley M, Ginter S. Risk Factors for Falls among Elderly Persons Living in the Community. *N Engl J Med*. 1988;319(26):1701-1707.

3. Friedman SM, Menzies IB, Bukata S V., Mendelson DA, Kates SL. Dementia and Hip Fractures: Development of a Pathogenic Framework for Understanding and Studying Risk. *Geriatr Orthop Surg Rehabil*. 2010;1(2):52-62.

4. Fernando E, Fraser M, Hendriksen J, Kim C, Hunter S. Risk factors for falls in peole with dementia: a systematic review. *Physiother Canada*. 2016;69(2):161-170.

5. Nichols E, Szoeke CEI, Vollset SE, et al. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol*. 2019;18(1):88-106.

6. Middleton A, Fritz SL. Assessment of Gait, Balance, and Mobility in Older Adults: Considerations for Clinicians. *Curr Transl Geriatr Exp Gerontol Rep*. 2013;2(4):205-214.

7. Freiberger E, De vreede P, Schoene D, et al. Performance-based physical function in older community-dwelling persons: A systematic review of instruments. *Age Ageing*. 2012;41(6):712-721.

8. Muir-Hunter SSWSSW, Graham L, Odasso MM, Montero-Odasso M, 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. *Physiother Canada*. 2015;67(3):255-262.

9. Bossers WW, Woude V Der, van der Woude L, Boersma F, Scherder E, van Heuvelen M. Recommended Measures for the Assessment of Cognitive and Physical Performance in Older Patients with Dementia: A Systematic Review. *Dement Geriatr Cogn Dis Extra*. 2012;2(1):589-609.

10. Trautwein S, Maurus P, Barisch-Fritz B, Hadzic A, Woll A. Recommended motor assessments based on psychometric properties in individuals with dementia: A systematic review. *Eur Rev Aging Phys Act*. 2019;16(1).

11. Van Ooteghem K, Musselman K, Gold D, et al. Evaluating mobility in advanced dementia: a scoping review and feasibility analysis. *Gerontologist*. 2018:1-14.

12. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *J Clin Epidemiol*. 2009;62(10):1006-1012.

13. National Institutes of Health. Quality assessment tool for observational cohart and cross-

sectional studies. Internet.

14. Lucas N, Macaskill P, Irwig P, et al. The reliability of a quality appraisal tool for studies of diagnostic reliability (QAREL). *BMC Med Res Methodol*. 2013;31(1):111.

15. Lucas NP, Macaskill P, Irwig L, Bogduk N. The development of a quality appraisal tool for studies of diagnostic reliability (QAREL). *J Clin Epidemiol*. 2010;63(8):854-861.

16. Sibley K, Beauchamp M, Van Ooteghem K, Straus S, Jaglal S. Using the Systems Framework for Postural Control to Analyze the Components of Balance Evaluated in Standardized Balance Measures: A Scoping Review. *Arch Phys Med Rehabil*. 2015;96(1):122-132.e29.

17. Horak FBF, Aspects P. Postural orientation and equilibrium: what do we need to know about neural control of balance to prevent falls? *Age Ageing*. 2006;35(suppl_2):ii7-ii11.

18. Streiner D, Norman G. *Health Measurement Scales: A Practical Guide to Their Development and Use.* New York: Oxfor University Press; 2003.

19. Cohen J. Statistical power analysis for the behavioral sciences. *Stat Power Anal Behav Sci*. 1988.

20. Tappen R, Roach K, Buchner D, Barry C, Edelstein J. Reliability of physical performance measures in nursing home residents with Alzheimer's disease. *J Gerontol A Biol Sci Med Sci*. 1997;52(1):M52-5.

21. Fox B, Henwood T, Neville C, et al. Relative and absolute reliability of functional performance measures for adults with dementia living in residential aged care. *Int Psychogeriatrics*. 2014;26(10):1659-1667.

22. Telenius EWE, Engedal K, Bergland A. Inter-rater reliability of the Berg Balance Scale, 30 s chair stand test and 6 m walking test, and construct validity of the Berg Balance Scale in nursing home residents with mild-to-moderate dementia. *BMJ Open*. 2015;5(9):1-7.

23. Lee HSH, Park SSW, Chung HKH. The Korean version of relative and absolute reliability of gait and balance assessment tools for patients with dementia in day care center and nursing home. *J Phys Ther Sci*. 2017;29(11):1934-1939.

24. Ries J, Echternach J, Nof L, Blodgett M. Test-Retest Reliability and Minimal Detectable Change Score for the Timed ' Up & Go" Test, the Six-Minute Walk Test, and Gait Spped In People with Alzheimer Disease. *Phys Ther*. 2009;89(6):569-579.

25. Blankevoort CG, Van Heuvelen MJG, Scherder EJA, C.G. B, M.J. van H. Reliability of Six Physical Performance Tests in Older People With Dementia. *Phys Ther*. 2013;93(1):69-78.

26. Lemke NNC, Wiloth S, Werner C, Hauer K. Validity, test-retest reliability, sensitivity to change and feasibility of motor-cognitive dual task assessments in patients with dementia.

Arch Gerontol Geriatr. 2017;70:169-179.

27. Olsen CCF, Bergland A, C.F. O, Bergland A. AO - Olsen CFO http://orcid. org/000.-0002-2841-9677. "Reliability of the Norwegian version of the short physical performance battery in older people with and without dementia." *BMC Geriatr*. 2017;17(1):1-10.

28. Sterke CS, van Beeck EF, Looman CWN, et al. An electronic walkway can predict short-term fall risk in nursing home residents with dementia. *Gait Posture*. 2012;36(1):95-101.

29. Sourdet S, Abellan Van Kan G, Soto ME, et al. Prognosis of an Abnormal One-Leg Balance in Community-Dwelling Patients With Alzheimer's Disease: A 2-Year Prospective Study in 686 Patients of the REAL.FR Study. *J Am Med Dir Assoc*. 2012;13(4):407.e1-407.e6.

30. Parfitt G, Post D, Penington A, Davison K, Corlis M. Challenges associated with physical assessments for people living with dementia: Modifying standard assessment protocols. *SAGE open Med*. 2020;8:2050312120910359.

31. Thomas VVS, Hageman PA, V.S. T, Thomas VVS, Hageman PA. A preliminary study on the reliability of physical performance measures in older day-care center clients with dementia. *Int Psychogeriatrics*. 2002;14(1):17-23.

32. Lee HS, Ko M, Park SW, Braden H. Concurrent validity of the Groningen Meander Walking and Timed Up and Go tests in older adults with dementia. *Physiother Theory Pract*. 2019;00(00):1-6.

33. 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. *Exp Gerontol*. 2019;115:9-18.

34. Lorbach E, Webster K, Menz H, et al. Physiological falls risk assessment in older people with Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2007;24(4):260-265.

35. Wittwer JJEJJE, Webster KEK, Andrews PPTPPT, et al. Test-retest reliability of spatial and temporal gait parameters of people with Alzheimer's disease. *Gait Posture*. 2008;28(3):392-396.

36. Suttanon P, Hill KDKKD, Dodd KKJK, Said CCM. Retest reliability of balance and mobility measurements in people with mild to moderate Alzheimer's disease. *Int Psychogeriatrics*. 2011;23(7):1152-1159.

37. Wittwer JJE, Webster KKE, Hill K, J.E. W, K.E. W. Reproducibility of gait variability measures in people with Alzheimer's disease. *Gait Posture*. 2013;38(3):507-510.

38. Sterke CSC, Huisman SSL, van Beeck EEF, et al. Is the Tinetti Performance Oriented Mobility Assessment (POMA) a feasible and valid predictor of short-term fall risk in nursing home residents with dementia? *Int Psychogeriatrics*. 2010;22(2):254-263.

39. McGough EL, Logsdon RGR, Kelly VEV, et al. Functional mobility limitations and falls in assisted living residents with dementia : Physical performance assessment and quantitative gait analysis. *J Geriatr Phys Ther*. 2013;36(2):78-86.

40. Bossers WJRWJR, van der Woude LLHV V., Boersma F, et al. The Groningen Meander Walking Test: A Dynamic Walking Test for Older Adults With Dementia. *Phys Ther*. 2014;94(2):262-272.

41. Panel for Prevention of Falls in Older Adults, American Geriatrics Society, Society BG, American Geriatrics Society and British Geriatrics Society. Summary of the Updated American Geriatrics Society/British Geriatrics Society Clinical Practice Guideline for Prevention of Falls in Older Persons. *J Am Geriatr Soc*. 2011;59(1):148-157.

42. Burton E, Cavalheri V, Adams R, et al. Effectiveness of exercise programs to reduce falls in older people with dementia living in the community: a systematic review and meta-analysis. *Clin Interv Aging*. 2015;10:421-434.

FIGURE CAPTION

Figure 1: Flow Diagram of Literature Search

Figure 1.

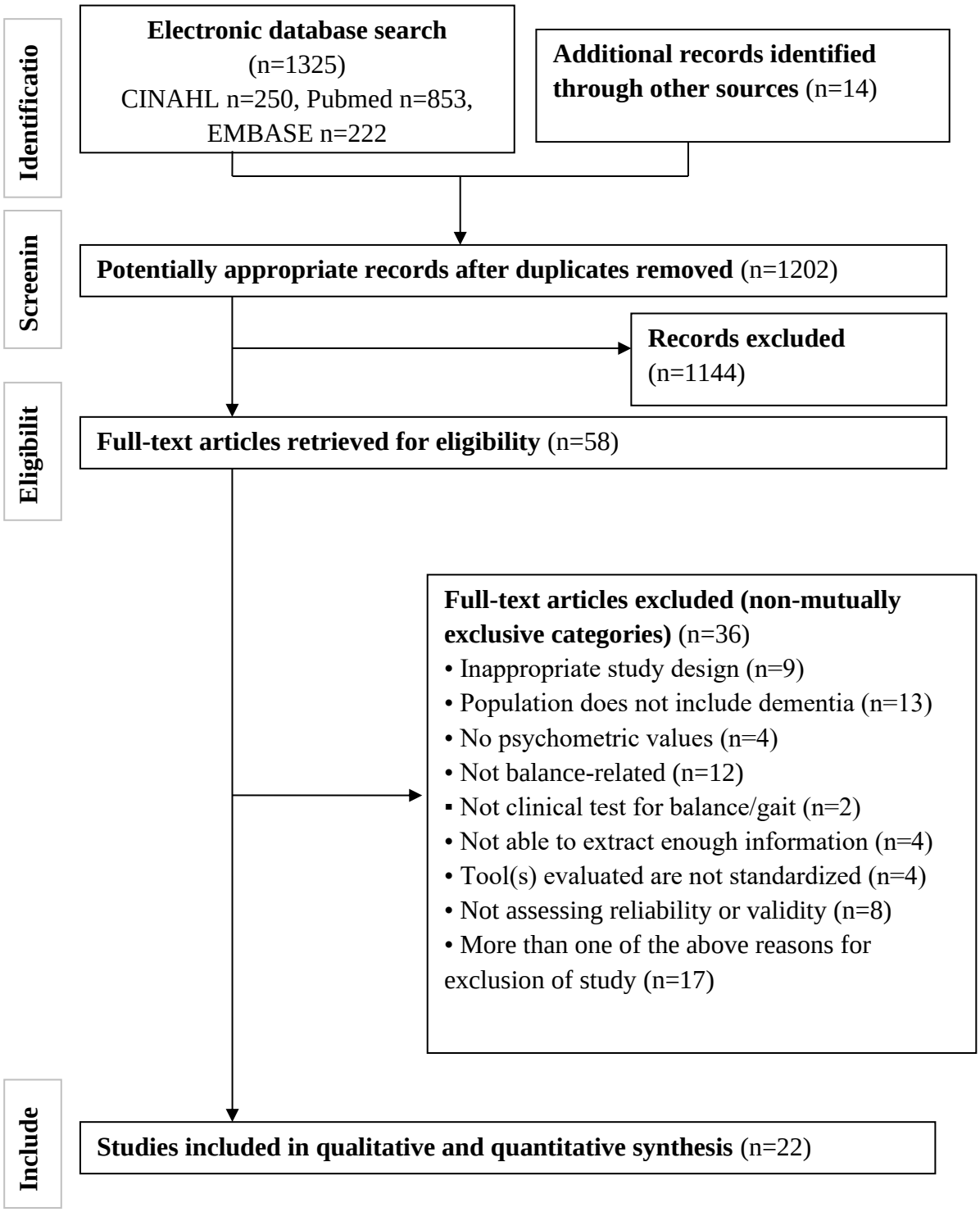
Identificatio
Electronic database search (n=1325)
CINAHL n=250, Pubmed n=853, EMBASE n=222
Additional records identified through other sources (n=14)
Screenin
Potentially appropriate records after duplicates removed (n=1202)
Records excluded (n=1144)
Eligibilit
Full-text articles retrieved for eligibility (n=58)
Full-text articles excluded (non-mutually exclusive categories) (n=36)
• Inappropriate study design (n=9)
• Population does not include dementia (n=13)
• No psychometric values (n=4)
• Not balance-related (n=12)
▪ Not clinical test for balance/gait (n=2)
• Not able to extract enough information (n=4)
• Tool(s) evaluated are not standardized (n=4)
• Not assessing reliability or validity (n=8)
• More than one of the above reasons for exclusion of study (n=17)
Include
Studies included in qualitative and quantitative synthesis (n=22)

Table 1: Summary of articles included in the systematic review

Author	Study Design (Test interval)	Study Objective	Mean Age (Age range) (years)	Sample Size (% female)	Setting	Dementia Diagnosis	Diagnostic Criteria for Dementia Diagnosis
Tappen et al. 1997[20] USA	Test-retest (1 week apart)	Examine the reliability of several established measures of physical performance in subjects with AD.	84.7 (79-92)	33 (65%)	Institution (Nursing Home)	Moderate to Severe AD	NINCDS-ADRDA
Thomas et al. 2002[31] USA	Test-retest (6-9 days apart)	Determine the test-retest reliability of a battery of established performance-based measures of strength and function among subjects with dementia.	80.5 ± 6.2 (NR)	12 (100%)	Community	Dementia	NR
Lorbach et al. 2007[34] Australia	Test-retest (1 week apart)	Determine the feasibility and test-retest reliability of a physiological test	AD: 79.3 ± 6.3 (NR) Control: 79.2 ± 6.0	Control: 21(53%) AD: 21(53%)	Community	Mild to Moderate AD	NINCDS-ADRDA & MMSE (scores of 11-26)

		battery designed to assess falls risk in people with AD.	(NR)				
Wittwer et al. 2008[35] Australia	Test-retest (1 week apart)	Examine test-retest reliability of spatiotemporal gait measurements in community dwelling people with AD.	80.6 ± 5.2 (70-91)	20 (10%)	Community	Mild to Moderate AD	NINCDS-ADRDA & MMSE (scores of 13-27)
Ries et al. 2009[24] USA	Test-retest (same day with a 30-60 minute rest interval)	Assess test-retest reliability and calculate MDCs for the TUG, 6MWT, and gait speed.	80.7 ± 8.8 (NR)	51 (66.7%)	Mixed: Community & Institution (Inpatient)	Mild to Severe AD	Information provided by facility/guardian & FAST (scores of 4-7)
Sterke et al. 2010[38] The Netherlands	Prospective	Evaluate the feasibility and examine the inter-rater reliability and the predictive ability of the POMA to predict fall risk in a population of nursing home residents with moderate to severe dementia.	81.0 ± 8 (NR)	75 (64%)	Institution (Nursing Home)	Moderate to Severe Dementia	DSM-IV-TR & GDS (stage 5 or 6)

Suttanon et al. 2011[36] Australia	Test-retest (1 week apart)	Determine safety, feasibility and retest reliability of clinical and force plate balance/mobility measurements in people with AD.	79.6 ± 6.2 (NR)	14 (50%)	Community	Mild to Moderate AD	MMSE (score of ≥ 10; confirmed diagnosis via medical practitioner)
Sourdet et al. 2012[29] France	Prospective	Explore the predictive value of an abnormal one leg balance test for function decline, nursing home admission, and mortality in community dwelling patients with AD.	(NR)	632 (71%)	Community	Mild to Moderate AD	NINCDS-ADRDA & MMSE (scores of 10-24)
Sterke et al. 2012[28] The Netherlands	Prospective	Evaluate the feasibility and validity of gait parameters measured with an electronic walkway system in predicting short-term fall risk in nursing home residents with dementia.	81.7 ± 7.0 (NR)	57 (61.2%)	Institution (Nursing home)	Moderate to Severe AD (133, 75.6%) or Unknown Dementia type (27, 15.3%)	DSM-IV-TR & GDS (stage 5 or 6)

Blankevoort et al. 2013[25] The Netherlands	Test-retest (1 week apart)	Determine the relative and absolute test-retest reliability of the 6MWT, figure 8 walk test, the TUG, the frailty and injuries cooperative studies of intervention techniques 4 balance test, chair rise test.	NR (70-92)	58 (70%)	Mixed: Community & Institution (Nursing Home)	Mild to Moderate Dementia	DSM-IV-TR & MMSE (scores of 10-28)
McGough et al. 2013[39] USA	Test-retest (4 hours)	Examine the test-retest reliability and concurrent validity of physical performance assessments and spatiotemporal gait measures in older adults with advanced dementia and to prospectively examine their relationship to functional mobility limitations and	83.6 ± 7.0 (67.4-95.6)	31 (93.5%)	Institution (Assisted Living)	Moderate to Severe Dementia	MMSE (scores of 2-26)

		falls over a 4-month period.					
Wittwer et al. 2013[37] Australia	Test-retest (1 week apart)	Investigate reproducibility of gait spatiotemporal variability measures in people with AD.	81.1 ± 5.2 (70.1-91.2)	16 (62.5%)	Community	Mild to Moderate AD	NINCDS-ADRDA & MMSE (scores of 13-26)
Bossers et al. 2014[40] The Netherlands	Test-retest (1 week apart)	Investigate the feasibility, test-retest reliability and MDC of GMWT.	86.7 ± 5.2 (NR)	42 (78.6%)	Institution (Nursing Home)	AD, vascular, Lewy body, AD & VD	Dementia diagnosed by a psychiatrist or medical doctor & MMSE (scores of 9-24)
Fox et al. 2014[21] Australia	Test-retest (1 week apart)	Assess relative and absolute test-retest reliability of commonly used functional performance measures in older adults with dementia residing in residential care facilities.	83.2 ± 9.9 (NR)	12 (91.7%)	Institution (Nursing Home)	Dementia	NR
Muir-Hunter et al. 2015[8] Canada	Test-retest (1 week apart)	Measure test-retest and interrater reliability of the BBS in	80.2 (NR)	15 (NR)	Community	Mild to Moderate AD	NINCDS-ADRDA by geriatrician & MMSE (score of 10-24)

		community-dwelling adults with mild to moderate AD.					
Telenius et al. 2015[22] Norway	Cross-sectional	Assess the inter-rater reliability between pairs of observers in the BBS, 30 second Chair Stand and 6mWT. Also investigate the internal consistency of the BBS.	82.7 ± 7.2 (66-91)	33 (75.8%)	Institution (Nursing Home)	Mild to Moderate Dementia	CDR (score of 1 or 2)
Lee et al. 2017[23] Korea	Test-retest (3-7 days apart)	Determine the relative and absolute reliability of Korean tools of the BBS, TUG, 4MWT, and GMWT in patients with dementia.	83.8 (NR)	53 (NR)	Mixed: Community & Institution (Nursing Home)	Dementia	Information provided by facility
Lemke et al. 2017[26] Germany	Test-retest (2-5 days apart)	Investigate validity, test-retest reliability, sensitivity to change, and feasibility of dual task assessments	82.7 ± 5.9 (NR)	105 (72.4%)	Mixed: Community & Institution (Nursing Home)	Mild to Severe Dementia	CERAD

		in patients with dementia.					
Olsen et al. 2017[27] Norway	Test-retest (1-7 days apart)	Establish test-retest reliability of the Norwegian version of the SPPB.	Control: 88.4 ± 9.2 (67-102) AD: 88.3± 6.2 (69-97)	Control: 37(78.4%) AD: 24(87.5%)	Mixed: Community & Institution (Nursing Home)	Dementia	Comprehensive geriatric assessment completed by physician
Chan et al. 2019[33] China	Test-retest (1-14 days apart)	Investigate the test-retest reliability, inter-rater reliability, minimal detectable change, and construct and known-group validity of the 2MWT, 6MWT and 10mWT using a progressive cueing system in frail older adults with dementia.	87.1 ± 6.2 (NR)	39 (92.3%)	Mixed: Community & Institution (Nursing Home)	Dementia	NR
Lee et al. 2019[32] South Korea	Cross-sectional	Determine the concurrent validity of the GMWT with the TUG in older adults with dementia.	83.4 ± 11.1 (NR)	57 (70.2%)	Institution (Nursing Home)	Dementia	By medical doctor

Parfitt et al. 2020[30] Australia	Test-retest (1 week apart)	Assess the test-retest reliability of the TUG, 2MWT and 4mWT using a staged cueing system in older adults living with dementia.	84.5 ± NR (69–94)	14 (100%)	Institution (Nursing Home)	Dementia	NR

Footnote: AD, Alzheimer's Disease; BBS, Berg Balance Scale; CERAD, Consortium to Establish a Registry for Alzheimer's Disease; CDR, Clinical Dementia Rating Scale; DSM-IV-TR, Diagnostic and Statistical Manual of Mental Disorders (fourth edition); FAST, Functional Assessment Staging Scale; GDS, Global Deterioration Scale; GMWT, Groningen Meander Walking Test; MDC, Minimal Detectable Change; MMSE, Mini-Mental State Examination; NINCDS-ADRDA, National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association; NR, Not Reported; POMA, Tinetti Performance Oriented Mobility Assessment; SPPB, Short Physical Performance Battery; TUG, Timed Up and Go Test; VD, Vascular Dementia; 2MWT, 2-Minute Walk Test; 4mWT, 4-Meter Walk Test; 6MWT, 6-Minute Walk Test; 6mWT, 6-Meter Walk Test; 10mWT, 10-Meter Walk Test.

Table 2: Summary of the clinical balance and gait tests evaluated among studies included in the review and the methodological reporting quality scores for each study.

Author	Tools Assessed	Psychometrics Measured	Quality Score	
			S1 (QAREL)	S2 (QOCCSS)
Tappen et al. 1997[20]	6 Minute Walk Test 25-Foot Walk Test (time to complete at usual pace)	Inter-rater, intra-rater reliability (ICC)	7/11	
Thomas et al. 2002[31]	Timed Up and Go Test 6-Meter Walk Test (time to complete at usual and fast pace)	Test-retest (ICC)	3/11	
Lorbach et al. 2007[34]	Physiological Profile Assessment	Test-retest (ICC)	4/11	
Wittwer et al. 2008[35]	Spatiotemporal Gait Parameters – usual pace (velocity, cadence, step length, stride length, support base, toe/out, swing time, stance time)	Test-retest (ICC), absolute reliability (MDC95)	4/11	
Ries et al. 2009[24]	Timed Up and Go Test 6 Minute Walk Test 15-Foot Walk Test (gait velocity at usual speed)	Test-retest (ICC), absolute reliability (SEM, MDC90)	5/11	
Sterke et al. 2010[38]	Tinetti Performance Oriented Mobility Assessment (Total score, balance and gait sub-scores)	Inter-rater reliability (ICC), predictive validity (area under the curve, sensitivity, specificity, positive predictive value, negative predictive value, hazard ratio)	5/11	10/14
Suttanon et al. 2011[36]	Modified Clinical Test for Sensory Interaction in Balance Functional Reach Test	Test-retest (ICC), absolute reliability (SEM, MDC95)	4/11	

	Step Test Limits of Stability Test Timed Up and Go Test - Motor Dual Task (while carrying a cup of water) Timed Up and Go Test - Cognitive Dual Task (while counting backwards by 3's) Quick Turn Test Spatiotemporal gait parameters – usual pace (velocity, step width, step length)			
Sourdet et al. 2012[29]	One Leg Balance Test	Predictive validity (hazard ratio)		11/14
Sterke et al. 2012[28]	Spatiotemporal Gait Parameters (velocity, mean stride length, heel-to-heel base of support variability, and double support time variability)	Predictive validity (area under the curve, sensitivity, specificity, positive predictive value, negative predictive value, odds ratios)		10/14
Blankevoort et al. 2013[25]	6-Meter Walk Test Figure 8 Walk Test Timed Up and Go Test Frailty and Injuries: Cooperative Studies of Intervention Techniques-4	Test-retest (ICC), absolute reliability (SEM, MDC95)	7/11	
McGough et al. 2013[39]	Spatiotemporal gait parameters – usual pace (velocity, cadence, swing time, stride length)	Test-retest reliability (ICC), concurrent validity (bivariate and partial correlations)	5/11	9/14
Wittwer et al. 2013[37]	Spatiotemporal Gait Parameters – usual pace (velocity, stride length, cadence, stride width, stride width variability)	Test-retest (ICC), absolute reliability (SEM, MDC95)	5/11	
Bossers et al. 2014[40]	Groningen Meander Walking Test	Test-retest reliability (ICC), absolute reliability (MDC95)	5/11	

Fox et al. 2014[21]	Short Physical Performance Battery Balance Outcome Measure for Elder Rehabilitation	Test-retest reliability (ICC)	4/11	
Muir-Hunter et al. 2015[8]	Berg Balance Scale Timed Up and Go Test Functional Reach Test	Test-retest, inter-rater reliability (ICC), absolute reliability (SEM, MDC95),	7/11	
Telenius et al. 2015[22]	Berg Balance Scale 6-Meter Walk Test (gait velocity at usual speed)	Inter-rater reliability (ICC), absolute reliability (SEM, MDC95)	5/11	
Lee et al. 2017[23]	Berg Balance Scale Timed Up and Go Test 4-Meter Walking Test (time to complete at usual speed) Groningen Meander Walking Test	Test-retest, inter-rater, (ICC), absolute reliability (SEM, MDC95)	6/11	
Lemke et al. 2017[26]	Dual Task Gait Test - walking and i) serial subtractions by 2 and ii) reciting alphabet) – gait velocity at usual pace	Test-retest reliability (ICC)	6/11	
Olsen et al. 2017[27]	Short Physical Performance Battery	Test-retest (ICC), absolute reliability (SEM, MDC95)	4/11	
Chan et al. 2019[33]	2-Minute Walk Test 6-Minute Walk Test 10-Meter Walk Test (gait velocity during the first 10 meters of the 2-Minute Walk Test and 6-Minute Walk Test)	Test-retest, inter-rater (ICC), absolute reliability (SEM, MDC95), concurrent validity (Spearman correlations)	7/11	8/14
Lee et al. 2019[32]	Groningen Meander Walking Test	concurrent validity (Spearman correlation)		5/14
Parfitt et al. 2020[30]	Timed Up and Go Test 2-Minute Walk Test 4-Meter Walking Test	Test-retest reliability (ICC)	4/11	

Footnote: ICC, Intraclass correlation coefficient; QAREL, Quality Appraisal Tool for Studies of Diagnostic Reliability; QOCCSS, Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies; MDC, Minimal Detectable Change; MDC90, Minimal Detectable Change 90% confidence; MDC95, Minimal Detectable Change 95% confidence; SEM, Standard Error of Measurement.

Table 3: Balance components assessed in standardized clinical tests of balance.

Tool	Anticipatory postural control	Cognitive processing	Dynamic stability	Functional stability limits	Motor system	Reactive postural control	Sensory strategies	Static stability	Verticality/ orientation in space	Total components assessed (#)
BBS (English,[8] Norwegian,[22] Korean[23])	✓		✓	✓	✓		✓	✓		6
BOOMER[21]	✓		✓		✓			✓		4
FISCIT-4[25]	✓				✓			✓		3
Functional Reach[8,36]	✓			✓	✓					3
Limits of Stability[36]	✓		✓	✓	✓			✓		5
mCTSIB[36]					✓		✓	✓		3
OLB[29]	✓				✓		✓	✓		4
POMA[38]	✓		✓		✓	✓	✓	✓		6
Step Test[36]	✓		✓		✓	✓				4
TUG (English[8,24,25,30,31] and Korean[23])	✓		✓		✓					3
TUG-cognitive[36]	✓	✓	✓		✓					4
TUG-motor[36]	✓	✓	✓		✓					4

Footnote: BBS, Berg Balance Scale; BOOMER, Balance Outcome Measure for Elder Rehabilitation; FISCIT-4, Frailty and Injuries: Cooperative Studies of Intervention Techniques; mCTSIB, Modified Clinical Test of Sensory Interaction on Balance; OLB, One Leg Balance Test; POMA, Tinetti Performance Oriented Mobility Assessment; TUG, Timed Up and Go.

Table 4: Recommended clinical tests for balance and gait assessment in people with dementia.

Test	Type of Dementia	Severity of Dementia
A. Community-dwelling		
Modified Test for Sensory Integration in Balance	Alzheimer's Disease	Mild to moderate
Spatiotemporal gait parameters	Alzheimer's Disease	Mild to moderate
Step Test	Alzheimer's Disease	Mild to moderate
Timed Up & Go Test	Alzheimer's Disease	Mild to moderate
B. Institution-dwelling		
Berg Balance Scale	Mixed (types not specified)	Mild to moderate
Groningen Meander Walk Test (*)	Alzheimer's Disease Vascular Dementia Mixed (Alzheimer's & Vascular) Lewy Body Dementia	Mild to moderate
6-meter Walk Test	Alzheimer's Disease	Moderate to severe
C. Mixed-dwelling		
Berg Balance Scale (Korean)	Mixed (types not specified)	Mild to moderate
Groningen Meander Walk Test (Korean)	Mixed (types not specified)	Mild to moderate
Timed Up & Go Test (English, Korean, Norwegian)	Mixed (types not specified)	Mild to moderate
2-Minute Walk Test (cued) (*)	Mixed (types not specified)	Not specified

6-Minute Walk Test (cued) (*)	Mixed (types not specified)	Not specified
10-Meter Walk Test (cued) (*)	Mixed (types not specified)	Not specified

Footnote: Recommended clinical tests of balance and gait with established validity are depicted by (*).

SUPPLEMENTARY DATA:

Supplementary Table 1: Sample of search strategy used for CINAHL database.

Database = CINAHL

Dementia: "dementia" OR "cognitive disorders" OR "delirium" OR "amnestic"

AND

Balance: "balance" OR "postural" OR "gait" OR "psychomotor performance"

AND

Psychometrics: "predictive value of tests" OR "reproducibility of results" OR "test re-test reliability" OR "instrument validation" OR "reliability" OR "validity" OR "pilot studies"

Supplementary Table 2: List of full-text articles reviewed and excluded with the reasons for exclusion from this systematic review.

Article Full Citation	Number of Exclusion Criteria Met	Exclusion Criteria Met
Beauchet O, Freiberger E, Annweiler C, Kressig R, Herrmann F, Allali G. Test-retest reliability of stride time variability while dual tasking in healthy and demented adults with frontotemporal degeneration. J. Neuroeng. Rehabil. 2011;8:37.	1	Not clinical test of gait (shoe insole foot switches)
Bossers, W. J., Van der Woude, L. H., Boersma, F., Scherder, E. J., & Van Heuvelen, M. J. Recommended measures for the assessment of cognitive and physical performance in older patients with dementia: a systemic review. Dementia and Geriatric Cognitive Disorders. 2012;2:589-609.	1	Study is a systematic review
Bramell-Risberg E, Jarnlo GB, Elmstahl S. Older women with dementia can perform fast alternating forearm movements and performance is correlated with tests of lower extremity function. Clin Inter in Aging. 2013;8:175-184.	1	Not balance/gait-related
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.	1	No diagnosis of dementia
Douglas A, von Kampen B, McAiney C, Shelley Wright. Adapting the tinetti tool for balance and gait for persons with dementia. Alzheimers Dement. 2015; 11(7): 234.	1	Not able to extract enough information
Farrell MK, Rutt RA, Lusardi MM, Williams AK. Are scores on the physical performance test useful in determination of risk of future falls in individuals with dementia?. J Geriatr Phys Ther. 2011;34(2):57-63.	1	Not balance/gait-related
Farrell MK, Rutt RA, Lusardi MM, Williams AK. Reliability of the physical performance test in people with Dementia. Phys. Occup. Ther. Geriatr. 2010;28:144–53.	1	Not balance/gait-related
Fox B, Henwood T, Neville C. Reliability of functional performance in older people with dementia. Australas J Ageing. 2013;32(4):248-9.	2	Not a research study; not able to extract enough information
Gonçalves J, Ansai JH, Masse FA, Vale FA, de Medeiros Takahashi AC, de Andrade LP. Dual-task as a predictor of falls in older people with mild	2	Not a standardized test; not able to extract enough information

cognitive impairment and mild Alzheimer's disease: a prospective cohort study. Brazilian journal of physical therapy. 2018; 22(5):417-23.		
Goodgold S, Kiami S, Ule D, Schoenberg A, Forman G. Applicability of the function reach and timed up and go tests for elderly individuals with alzheimer's disease: pilot investigation. Phys & Occp Ther in Geriatr. 2001;19(1):21-36.	2	No psychometric values; not assessing reliability or validity
Hauer, K. A., Kempen, G. I., Schwenk, M., Yardley, L., Beyer, N., Todd, C., Oster, P., Zijlstra, G. R. Validity and sensitivity to change of the falls efficacy scales international to assess fear of falling in older adults with and without cognitive impairment. Gerontology. 2011;57:462-472.	2	Not testing physical performance in relation to balance; no diagnosis of dementia
Hauer, K. Measuring functional performance in persons with Dementia. Journal of the American Geriatrics Society. 2008;56(5):949-950.	4	Not a research study; not able to extract enough information; no psychometric values; not assessing reliability or validity
Hauer, K., Yardley, L., Beyer, N., Kempen, G., Dias, N., Campbell, M., Becker, C., Todd, C. Validation of the Falls Efficacy Scale and Falls Efficacy Scale International in geriatric patients with and without cognitive impairment: results of self-report and interview-based questionnaires. Gerontology. 2010;56:190-199.	2	Not testing physical performance in relation to balance; no diagnosis of dementia
Hesseberg K, Bentzen H, Bergland A. Reliability of the senior fitness test in community-dwelling older people with cognitive impairment. Physiother Res Int. 2015;20(1):37-44.	1	No diagnosis of dementia
Holbein-Jenny, M., Billek-Sawhney, B., Beckman, E., & Smith, T. Balance in personal care home residents: a comparison of the Berg Balance Scale, the Multi-Directional Reach Test, and the Activities-Specific Balance Confidence Scale. Journal of Geriatric Physical Therapy 2005;28(2):48-53.	1	No diagnosis of dementia
Kaur N, Belchior P, Gelinas I, Bier N. Critical appraisal of questionnaires to assess functional impairment in individuals with mild cognitive impairment. Int Psychogeriatr. 2016;28(9):1425-39.	3	Study is a systematic review; not balance-related; no diagnosis of dementia
Mahurin RK1, DeBettignies BH, Pirozzolo FJ. Structured assessment of independent living skills: preliminary report of a performance measure of functional abilities in dementia. J Gerontol. 1991 Mar;46(2):P58-66.	1	Not balance/gait-related

McGough EL, Lin S, Belza B, Becofsky KM, Jones DL, Lie M, et al. A scoping review of physical performance outcome measures used in exercise interventions for older adults with Alzheimer disease and related dementias. J Geriatr Phys Ther. 2017;00:1-20.	1	Study is a scoping review
McMichael KA, Vander Bilt J, Lavery L, Rodriguez E, Ganguli M. Simple balance and mobility tests can assess falls risk when cognition is impaired. Geriatr Nurs. 2008;29(5):311-23.	3	No psychometric values; no diagnosis of dementia; not assessing reliability or validity
Moe-Nilssen, R., Nordin, E., & Lundin-Olsson, L. Criteria for evaluation of measurement properties of clinical balance measures for use in fall prevention studies. Journal of Evaluation in Clinical Practice. 2008;14:236-240.	2	No diagnosis of dementia; not a standardized test
Montero-Odasso M, Casas A, Hansen KT, Bilski P, Gutmanis I, et al. Quantitative gait analysis under dual-task in older people with mild cognitive impairment: a reliability study. J Neuroeng Rehabil. 2009;6:35.	1	No diagnosis of dementia
Rockwood, K., Awalt, E., Carver, D., & MacKnight, C. Feasibility and measurement properties of the Functional Reach and the Timed Up and Go Tests in the Canadian Study of Health and Aging. Journal of Gerontology. 2000;55A(2):70-73.	1	No diagnosis of dementia
Rolenz, E., & Reneker, J. C. Validity of the 8-Foot Up and Go, Timed Up and Go, and Activities-Specific Balance Confidence scale in older adults with and without cognitive impairment. Journal of Rehabilitation Research and Development. 2016;53(4): 511-518.	2	No diagnosis of dementia; not assessing reliability or validity
Rolland Y, Abellan van Kan G, Nourhashemi F, Andrieu S, Cantet C, et al. An abnormal “one-leg balance” test predicts cognitive decline during Alzheimer’s disease. J Alzheimers Dis. 2009;16(3):525-31.	2	No psychometric values; not assessing reliability or validity
Ryan JJ, McCloy C, Rundquist P, Srinivasan V, Laird R. Fall risk assessment among older adults with mild alzheimer disease. J of Geriatr Phys Ther. 2011;34(1):19-27.	1	Not balance/gait-related
Scott, V., Votova, K., Scanlan, A., & Close, J. Multifactorial and functional mobility assessment tools for fall risk among older adults in community, home-support, long-term and acute care settings. Age and Aging. 2007;36:130-139.	2	Study is a systematic review; no diagnosis of dementia

Suzuki M, Kirimoto H, Inamura A, Yagi M, Omori Y, Yamada S. The relationship between knee extension strength and lower extremity functions in nursing home residents with dementia. Disability and rehabilitation. 2012 Feb 1;34(3):202-9.	1	Not assessing reliability or validity
Suzuki M, Yamada S, Inamura A, Omori Y, Kirimoto H, Sugimura S, Miyamoto M. Reliability and validity of measurements of knee extension strength obtained from nursing home residents with dementia. Am J Phys Med Rehabil. 2009;88(11):924-33.	1	Not balance/gait-related
Tappen RM. J Gerontol Nurs. Development of the refined ADL assessment scale. 1994;20(6):36-42.	2	Not a research study; not balance/gait-related
Thapa PB, Gideon P, Fought RL, Kormicki M, Ray WA. Comparison of clinical and biomechanical measures of balance and mobility in elderly nursing home residents. J Am Geriatr Soc. 1994;42(5):493-500.	1	No diagnosis of dementia
Thunborg, C., Von Heideken Wagert, P., Ivarsson, A. B., & Soderlund, A. Inter- and intra-rater reliability of a newly developed assessment scale: The Dyadic Interaction in Dementia Transfer Assessment Scale (DIDTAS). Physical and Occupational Therapy in Geriatrics. 2015;33(4):279-293.	2	Not testing physical performance in relation to balance; not a standardized test
Van Iersel, M. B., Benraad, C. E., & Olde Rikkert, M. G. Validity and reliability of quantitative gait analysis in geriatric patients with and without Dementia. Journal of the American Geriatrics Society. 2007;55(4):632-634.	2	Not a research study; not able to extract enough information
Van Ooteghem, K., Musselman, K., Gold, D., Marcil, M., Keren, R., Tartaglia, M. C., Flint, A., & Iaboni, A. Evaluating mobility in advanced Dementia: a scoping review and feasibility analysis. The Gerontologist. 2018;00:1-14.	1	Study is a systematic review
Vidoni ED1, Billinger SA, Lee C, Hamilton J, Burns JM. The Physical Performance Test predicts aerobic capacity sufficient for independence in early-stage Alzheimer disease. J Geriatr Phys Ther. 2012;35(2):72-8.	2	Not balance/gait-related; not assessing reliability or validity
Werner, C., Wiloth, S., Lemke, N. C., Kronbach, F., & Hauer, K. Development and validation of a novel motor-cognitive assessment strategy of compensatory sit-to-stand maneuvers in people with Dementia. Journal of Geriatric Physical Therapy. 2018;41:143-154.	2	Not a standardized test; not balance/gait-related
Wittwer JE, Webster KE, Hill K. The effects of a concurrent motor task on walking in Alzheimer's disease. Gait & posture. 2014 Jan 1;39(1):291-6.	1	Not assessing reliability or validity

Supplementary Table 3: Psychometric properties of standardized clinical tests of balance or gait evaluated in community-dwelling samples of older adults living with dementia.

	Relative Reliability			**Absolute Reliability**	
	Inter-rater Reliability	**Intra-rater Reliability**	**Test-retest**	**SEM**	**MDC_{95}**
Berg Balance Scale	0.72 (0.31-0.91)[8]		0.95 (0.85-0.98)[8]	6.01[8] (t-r)	16.66[8] (t-r)
Functional Reach (cm)	0.79 (0.43-0.94)[8]		0.84[36] 0.81 (0.52-0.94)[8]	1.61[36] 4.56[8] (t-r)	3.15[36] 12.64[8] (t-r)
Limits of Stability					
Directional Control (%)			0.71[36]	5.24[36]	10.27[36]
Maximum Excursion (%)			0.68[36]	4.44[36]	8.71[36]
Movement Velocity (degree/sec)			0.48[36]	0.46[36]	0.91[36]
mCTSIB			0.91[36]	0.17[36]	0.34[36]
Physiological Profile Assessment			0.69 (0.37-0.86)[34]		
Quick Turn Test					
Step Quick Turn - Sway (degrees)			0.64[36]	4.56[36]	8.93[36]
Step Quick Turn – Time (sec)			0.55[36]	0.33[36]	0.64[36]
Walk Across - Speed (cm)			0.50[36]	7.58[36]	14.86[36]
Walk Across - Step Length (cm)			0.75[36]	4.59[36]	9.0[36]
Walk Across - Step Width (cm/sec)			0.89[36]	1.26[36]	2.48[36]
Spatiotemporal Gait Parameters					
Cadence (steps/min)			0.88 (0.72-0.95)[35] 0.65 (0.25-0.86)[37]	0.8[37]	8.13[35] 2.3[37]
Stance Time, L (secs)			0.86 (0.68-0.94)[35]		0.07[35]
Stance Time, R (secs)			0.87 (0.70-0.95)[35]		0.06[35]
Step Length, L (cm)			0.96 (0.91-0.99)[35]		5.80[35]
Step Length, R (cm)			0.97 (0.93-0.99)[35]		5.27[35]
Stride Length (cm)			0.80 (0.52-0.93)[37]	1.1[37]	3.1[37]
Stride Length, L (cm)			0.97 (0.92-0.99)[35]		11.17[35]
Stride Length, R (cm)			0.97 (0.93-0.99)[35]		10.24[35]

Stride Width (cm)			0.83 (0.59-0.94)[37]	3.0[37]	8.3[37]
Support Base, L (cm)			0.92 (0.81-0.97)[35]		2.26[35]
Support Base, R (cm)			0.92 (0.82-0.97)[35]		2.23[35]
Swing Time, L (sec)			0.93 (0.84-0.97)[35]		0.03[35]
Swing Time, R (sec)			0.90 (0.76-0.96)[35]		0.03[35]
Toe in/out, L (degrees)			0.96 (0.94-0.99)[35]		3.09[35]
Toe in/out, R (degrees)			0.91 (0.78-0.96)[35]		3.06[35]
Velocity (m/s)			0.95 (0.88-0.98)[35] 0.66 (0.26-0.87)[37]	 1.6[37]	0.13[35] 4.4[37]
Step Test (reps)			0.87[36]	1.24[36]	2.42[36]
Timed Up & Go Test (sec)	0.98 (0.93-0.99)[8]		0.87[31] 0.76[36] 0.72 (0.33-0.90)[8]	 1.24[36] 1.24[8] (t-r)	 2.42[36] 3.44[8] (t-r)
Timed Up & Go Test – Cognitive Task					
TUG while Counting Backwards by 3's (time in secs)			0.51[36]	2.39[36]	4.69[36]
Timed Up & Go Test - Motor Task					
TUG while carrying a glass of water (time in secs)			0.70[36]	1.45[36]	2.83[36]
6-Meter Walk Test					
Steps, Usual Gait			0.80[31]		
Time, Usual Gait (sec)			0.92[31]		
Steps, Fast Gait			0.90[31]		
Time, Fast Gait (sec)			0.95[31]		

Footnote: mCTSIB, Modified Clinical Test of Sensory Interaction on Balance; MDC, Minimal Detectable Change; MDC95, Minimal Detectable Change 95% confidence; SEM, Standard Error of Measurement; t-r, Test-retest; TUG, Timed Up and Go.

Supplementary Table 4: Psychometric properties of standardized clinical tests of balance or gait evaluated in institution-dwelling samples of older adults living with dementia.

	Relative Reliability			**Absolute Reliability**	
	Inter-rater Reliability	**Intra-rater Reliability**	**Test-retest**	**SEM**	**MDC_{95}**
Berg Balance Scale	0.996[22]			0.97[22]	1.92[22]
BOOMER					
Functional Reach (cm)			0.38[21]		
Static Timed Standing (sec)			0.47[21]		
Step Test L (reps)			0.79[21]		
Step Test R (reps)			0.70[21]		
TUG (sec)			0.86[21]		
Groningen Meander Walk Test *(sec)*					
Overall			0.94 (0.90-0.97)[40]	1.93 (1.64-2.54)[40]	5.35[40]
Overstep (n)			0.63 (0.41-0.78)[40]	1.58 (1.31-2.03)[40]	4.38[40]
Overstep - No aid			0.67 (0.37-0.85)[40]	0.98 (0.77-1.41)[40]	2.71[40]
Overstep - With 4WW			0.58 (0.18-.81)[40]	2.09 (1.61-3.15)[40]	5.78[40]
With 4WW			0.75 (0.71-0.95)[40]	3.73 (2.02-3.97)[40]	10.35[40]
Without Walking Aid			0.97 (0.94-0.99)[40]	1.07 (0.84-1.54)[40]	2.96[40]
Short Physical Performance Battery					
SPPB Summary Score			0.88[21]		
Standing Balance (sec)			0.49[21]		
2.4-Meter Walk (sec)			0.68[21]		
5 Repeated Chair Stands (sec)			0.97[21]		
Spatiotemporal Gait Parameters					
Cadence (Steps/min)			0.91 (0.62-0.99)[39]		

Gait Speed (m/s)			0.95 (0.81-0.99)[39]		
Stride Length (cm)			0.97 (0.87-0.99)[39]		
Swing Time (sec)			0.96 (0.81-0.99)[39]		
Timed Up & Go Test (cued)			0.91[30] (ICC) 0.87[30] (r)		
Tinetti Performance Oriented Mobility Assessment					
Balance	0.97[38]				
Gait	0.88[38]				
Total	0.97[38]				
2-Minute Walk Test (cued)			0.87[30] (r)		
4-Meter Walk Test (cued)			0.84[30] (ICC) 0.83[30] (r)		
6-Meter Walk Test (m/s)	0.97[22]			0.03[22]	0.06[22]
6-Minute Walk Test					
Distance (feet)	0.97-0.99[20]	0.76-0.90[20]			
Speed (feet/sec)	0.96-0.98[20]	0.75-0.89[20]			
25-Foot Walk					
Distance (feet)	0.85-0.97[20]	0.57-0.73[20]			
Speed (feet/sec)	0.83-0.94[20]	0.45-0.77[20]			

Footnote: BOOMER, Balance Outcome Measure for Elder Rehabilitation; ICC, Intraclass correlation coefficient; MDC, Minimal Detectable Change; MDC95, Minimal Detectable Change 95% confidence; r, Pearson Correlation between test re-test assessments; SEM, Standard Error of Measurement; SPPB, Short Physical Performance Battery; t-r, Test-retest; TUG, Timed Up and Go; 4WW, 4-Wheeled Walker.

Supplementary Table 5: Psychometric properties of standardized clinical tests of balance or gait evaluated in mixed-dwelling samples of older adults living with dementia.

	Relative Reliability			**Absolute Reliability**	
	Inter-rater Reliability	**Intra-rater Reliability**	**Test-retest**	**SEM**	**MDC_{95}**
Berg Balance Scale (Korean Version)	0.99[23]		0.99[23]	0.78[23] (inter-r) 1.36[23] (t-r)	2.18[23] (inter-r) 3.78[23] (t-r)
Dual Task: Walking and Cognition (speed, cm/s)					
Reciting ABC's			0.96 (0.93-0.97)[26]		
Naming animals			0.96 (0.93-0.97)[26]		
Naming plants			0.95 (0.92-0.96)[26]		
Counting Forward by 2			0.93 (0.90-0.96)[26]		
Counting Backwards by 3			0.92 (0.87-0.95)[26]		
Figure of 8 Walk Test (sec)					
All Severities			0.91 (0.85-0.95)[25]	6.26 (5.41-8.21)[25]	17.35[25]
Mild			0.94 (0.86-0.97)[25]	6.24 (5.63-10.03)[25]	17.30[25]
Moderate			0.85 (0.67-0.94)[25]	6.00 (4.01-7.58)[25]	16.63[25]
FISCIT-4					
All Severities			0.79 (0.67-0.87)[25]	0.55 (0.47-0.69)[25]	1.52[25]
Mild			0.82 (0.65-0.91)[25]	0.59 (0.48-0.81)[25]	1.64[25]
Moderate			0.80 (0.61-0.90)[25]	0.60 (0.48-0.82)[25]	1.66[25]
Groningen Meander Walk Test (Korean version)					
Seconds	0.99[23]		0.99[23]	1.00[23] (inter-r) 1.36[23] (t-r)	2.78[23] (inter-r) 3.78[23] (t-r)
Step Count	0.99[23]		0.96[23]	0.76[23] (inter-r) 1.49[23] (t-r)	2.12[23] (inter-r) 4.13[23] (t-r)
Short Physical Performance Battery (Norwegian Version)					

Balance Subtest			0.74 (0.38-0.89)[27]	0.48[27]	1.33[27]
Sit to Stand Subtest			0.83 (0.72-0.90)[27]	0.21[27]	0.58[27]
SPPB Sum Score			0.84 (0.64-0.93)[27]	0.68[27]	1.88[27]
Walking Subtest			0.96 (0.92-0.98)[27]	0.19[27]	0.53[27]
4-Meter Walk Test (m/s)			0.94 (0.85-0.97)[27]	0.10[27]	0.28[27]
Spatiotemporal Gait Parameters					
Gait Speed (cm/s)			0.97-0.98[24]	5.72[24] (all severities) 6.07[24] (mild-moderate) 5.48[24] (moderate-severe)	9.44[24] (MDC90) (all severities)
Timed Up & Go Test (sec)					
All Severities			0.99-0.99[24] 0.94 (0.92-0.97)[25]	2.48[24] 2.12 (1.74-2.52)[25]	4.09[24] (MDC90) 5.88[25]
Mild			0.96 (0.92-0.98)[25]	1.43 (1.06-1.79)[25]	3.96[25]
Mild-Moderate				1.52[24]	
Moderate			0.94 (0.87-0.97)[25]	2.91 (2.10-3.61)[25]	8.07[25]
Moderate-Severe				3.03[24]	
Timed Up & Go Test (Korean version) (sec)	0.99[23]		0.99[23]	0.63[23] (inter-r) 1.27[23] (t-r)	1.75[23] (inter-r) 3.52[23] (t-r)
2-Minute Walk Test (cued) (m)	0.92 (0.86-0.96)[33] (visit #1) 0.96 (0.92-0.98)[33] (visit #2)		0.98 (0.96-0.99)[33]	3.3[33] (t-r)	9.1[33] (t-r)
4-Meter Walk Test (Korean version) (sec)	0.82[23]		0.85[23]	0.74[23] (inter-r) 0.64[23] (t-r)	2.06[23] (inter-r) 1.78[23] (t-r)
6-Meter Walk Test (m/s)					
All Severities			0.86 (0.78-0.92)[25]	0.10 (0.08-0.12)[25]	0.27[25]
Mild			0.83 (0.67-0.91)[25]	0.11 (0.09-0.11)[25]	0.29[25]
Moderate			0.89 (0.78-0.95)[25]	0.09 (0.07-0.13)[25]	0.25[25]
6-Minute Walk Test (feet)					
All Severities			0.98-0.99[24]	66.53[24]	109.8[24] (MDC90)
Mild-Moderate				71.72[24]	

Moderate-Severe				64.20[24]	
6-Minute Walk Test (cued) (m)	0.95 (0.91-0.97)[33] (visit #1) 0.94 (0.89-0.97)[33] (visit #2)		0.98 (0.97-0.99)[33] (t-r)	10.1[33] (t-r)	28.1[33] (t-r)
10-Meter Walk Test (cued) (m/s)	0.91 (0.83–0.95)[33] (2M: visit #1) 0.86 (0.75–0.93)[33] (2M: visit #2) 0.94 (0.89–0.97)[33] (6M: visit #1) 0.65 (0.42–0.80)[33] (6M: visit #2)		0.91 (0.83-0.95)[33] (2M) (t-r) 0.94 (0.89-0.97)[33] (6M) (t-r)	0.06[33] (2M) (t-r) 0.06[33] (6M) (t-r)	0.17[33] (2M) (t-r) 0.16[33] (6M) (t-r)

Footnote: FISCIT-4, Frailty and Injuries: Cooperative Studies of Intervention Techniques; inter-r, Inter-rater Reliability; MDC, Minimal Detectable Change; MDC90, Minimal Detectable Change 90% confidence; MDC95, Minimal Detectable Change 95% confidence; SEM, Standard Error of Measurement; SPPB, Short Physical Performance Battery; t-r, Test-retest; 2M, 10-Meter Walk Test measured during the 2-Minute Walk Test (cued); 6M, 10-Meter Walk Test measured during the 6-Minute Walk Test (cued).

Supplementary Table 6: Validity of standardized clinical tests of balance or gait evaluated in community and institutional samples of older adults living with dementia.

	Concurrent Validity	Predictive validity						
		AUC (95% CI)	Cut-off	Sn	Sp	PPV	NPV	Other
COMMUNITY SETTING								
One Leg Balance Test (< 5 seconds)								Functional Decline (adjusted HR): 1.69 (1.26-2.26)[29]
								Walking Abilities (adjusted HR): 2.56 (1.58-4.15)[29]
								Nursing Home Admission (adjusted HR): 2.51 (1.69-3.73)[29]
								Death (adjusted HR): 2.42 (1.43-4.11)[29]
INSTITUTIONAL SETTING								
Groningen Meander Walking Test (s)	TUG (ρ): 0.69[32]							
Spatiotemporal Gait Parameters								
Cadence (Steps/min)	Bivariate correlation (r) SPPB: 0.56, mBBS: 0.68[39]							
	Partial correlation (r, adjusted) SPPB: 0.51, mBBS: 0.67[39]							
Cycle Double Support (%)	Bivariate correlation (r) SPPB: -0.69, mBBS: -0.70[39]							
	Partial correlation (r, adjusted) SPPB: -0.64, mBBS: -0.64[39]							
Double Support Time Variability (CoV)		0.59 (0.50–0.68)[28]	9[28]	63%[28]	51%[28]	30%[28]	58%[28]	3-month Falls Occurrence (adjusted OR): 1.53 (1.05-2.25)[28] (10% increase)

Gait Speed (m/s)	Bivariate correlation (r) SPPB: 0.66, mBBS: 0.73[39]							
	Partial correlation (r, adjusted) SPPB: 0.71, mBBS: 0.62[39]							
Heel-to-heel Base of Support Variability (CoV)		0.59 (0.51–0.68)[28]	17[28]	60%[28]	56%[28]	44%[28]	70%[28]	3-month Falls Occurrence (adjusted OR): 1.49 (1.15-1.93)[28] (10% decrease)
Stride Length (cm)	Bivariate correlation (r) SPPB: 0.63, mBBS: 0.72[39]	0.67 (0.59-0.75)[28]	85[28]	86%[28]	52%[28]	50%[28]	84%[28]	3-month Falls Occurrence (adjusted OR): 1.19 (1.03-1.40)[28] (10cm decrease)
	Partial correlation (r, adjusted) SPPB: 0.61, mBBS: 0.71[39]							
Stride Length Variability (CoV)	Bivariate correlation (r) SPPB: -0.71, mBBS: -0.85[39]							
	Partial correlation (r, adjusted) SPPB: -0.65, mBBS: -0.54[39]							
Stride Velocity (m/s)	Bivariate correlation (r) SPPB: 0.65, mBBS: 0.72[39]							
	Partial correlation (r, adjusted) SPPB: 0.66, mBBS: 0.72[39]							
Swing Time Variability (CoV)	Bivariate correlation (r) SPPB: -0.71, mBBS: -0.76[39]							
	Partial correlation (r, adjusted) SPPB: -0.60, mBBS: -0.67[39]							
Velocity (cm/s)		0.66 (0.58-0.74)[28]	68[28]	82%[28]	52%[28]	49%[28]	82%[28]	3-month Falls Occurrence (adjusted OR): 1.22 (1.04-1.43)[28] (10cm/s decrease)
Tinetti Performance Oriented Mobility Assessment								
Balance		0.67 (0.52-0.81)[38]	≤11[38]	70%[38]	51%[38]	35%[38]	81%[38]	3-month Falls Occurrence (unadjusted HR): 1.11 (1.01-1.23)[38]

Gait		0.67 (0.53-0.81)[38]	≤9[38]	70%[38]	61%[38]	37%[38]	81%[38]	
Total		0.70 (0.53-0.81)[38]	≤21[38]	85%[38]	56%[38]	38%[38]	89%[38]	3-month Falls Occurrence (adjusted HR): 1.08 (1.00-1.17)[38]
MIXED SETTING								
Dual Task: Walking and Cognition (speed, cm/s)								
Reciting ABC's	TUG (r): -0.84[26]							
	POMA (r): 0.80[26]							
Naming animals	TUG (r): -0.82[26]							
	POMA (r): 0.73[26]							
	MMSE (r): 0.28[26]							
Naming plants	TUG (r): -0.78[26]							
	POMA (r): 0.69[26]							
	MMSE (r): 0.29[26]							
Counting Forward by 2	TUG (r): -0.79[26]							
	POMA (r): 0.70[26]							
Counting Backwards by 3	TUG (r): -0.74[26]							
	POMA (r): 0.62[26]							
	MMSE (r): 0.37[26]							
2-Minute Walk Test (cued) (m)	EMS (ρ): 0.43[33]							
	BBS (ρ): 0.49[33]							
	MBI (ρ): 0.54[33]							
6-Minute Walk Test (cued) (m)	EMS (ρ): 0.39[33]							
	BBS (ρ): 0.47[33]							

	MBI (ρ): 0.48[33]							
10-Meter Walk Test (cued) (m/s)	EMS (ρ; 2M): 0.39[33]							
	BBS (ρ; 2M): 0.49[33]							
	MBI (ρ; 2M): 0.45[33]							
	EMS (ρ; 6M): 0.27[33]							
	BBS (ρ; 6M): 0.35[33]							
	MBI (ρ; 2M): 0.46[33]							

Footnote: AUC, Area Under the Curve; BBS, Berg Balance Scale; CI, Confidence Interval; CoV, Coefficient of Variation; EMS, Elderly Mobility Scale; HR, Hazard Ratio; mBBS, Modified Berg Balance Scale; MBI, Modified Barthel Index; MMSE, Mini-Mental State Examination; NPV, Negative Predictive Value; OR, Odds Ratio; POMA, Tinetti Performance Oriented Mobility Assessment; PPV, Positive Predictive Value; Sn, Sensitivity; Sp, Specificity; SPPB, Short Physical Performance Battery; TUG, Timed Up and Go; 2M, 10-Meter Walk Test measured during the 2-Minute Walk Test (cued); 6M, 10-Meter Walk Test measured during the 6-Minute Walk Test (cued).