https://doi.org/10.1016/j.maturitas.2015.01.001

Title: Psychometric Properties of Dual-Task Balance Assessments for Older Adults: A Systematic Review

Authors:

L. YANG[a,b], L.R. LIAO[b], F.M.H. LAM[b], C.Q. HE[c], M.Y.C. PANG[b]*

Affiliations:

[a] Institute for Disaster Management and Reconstruction, Sichuan University – Hong Kong Polytechnic University, Sichuan, China

[b] Department of Rehabilitation Sciences, Hong Kong Polytechnic University, Hong Kong, China

[c] Department of Rehabilitation Medicine, West China Hospital, Sichuan University, Chengdu, Sichuan Province, China

*Corresponding Author. Tel: +852-2330-8656; Fax: +852-2766-7156

Email address: Marco.Pang@polyu.edu.hk (M.Y.C. Pang)

ABSTRACT

Background: The ability to maintain balance while simultaneously performing a cognitive task is essential for daily living and has been implicated as a risk factor of falls in older adults.

Aims: To evaluate the evidence related to the psychometric properties of dual-task balance assessments in older adults.

Methods: An extensive literature search of electronic databases was conducted. Articles were included if they evaluated the psychometric properties of dual-task balance assessment tools in older adults. The data were extracted by two independent researchers and confirmed with the principal investigator. The methodology quality of each study was rated by using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist.

Results: Twenty-six articles were included in this systematic review. For dual-task static standing balance assessments, the center of pressure-related parameters (displacement, velocity) and reaction time measurements were reliable but not useful for prediction of falls. For walking balance assessments, the gait outcomes derived generally demonstrated good to excellent reliability [Intraclass correlation coefficient >0.75)], but their ability to predict falls varied. Outcomes derived from the cognitive tasks and the dual-task cost (dual-task performance minus single-task performance) mostly demonstrated low to fair reliability. The methodological quality of majority of studies was poor to fair, mainly due to small sample size.

Conclusions: Among the dual-task balance assessments examined, the reliability and validity varied. The findings of this review should be useful in guiding the selection of dual-task balance measures in future research.

Keywords: systematic review; dual-task; psychometrics; balance; older adults

Word Count: 2994

1. INTRODUCTION

Deficits in balance and mobility are among the major health concerns in the elderly population [1-3]. Functional community ambulation not only necessitates a critical level of balance skills, but also the ability to engage in cognitive tasks while simultaneously performing the walking task (i.e., dual-tasking) in constantly changing environments [4]. Increasing research has examined the dual-task balance and mobility performance in older adults and its clinical correlates.

Different assessment tools have been developed to evaluate dual-task balance and mobility performance [5-13]. There is evidence that dual-task balance performance declines with aging, particularly when the task is more challenging in nature [14-16]. Impaired dual-task performance has also been associated with increased risk of falls in the elderly [17-19]. While various dual-task balance assessment tools are available, studying their psychometric properties is essential. For example, good reliability and validity are required for accurate evaluation of patient performance. Adequate responsiveness of a measurement tool is important for detecting change in dual-task performance over time and assessing treatment effectiveness [20].

Previous reviews have only evaluated the ability of various dual-task assessments to predict falls (i.e., predictive validity) [17-19] without addressing other important aspects of psychometric properties (reliability, convergent validity, etc.). Moreover, none of the reviews used a systematic review approach. To date, no systematic review has examined the psychometric properties of different dual-task balance assessment tools in older adults. The current systematic review was undertaken to address this knowledge gap.

2. METHODS

2.1. Search Strategy

An extensive literature search of electronic databases was conducted, including PubMed, CINAHL (1982–9 December 2013), MEDLINE (1950–9 December 2013), PsycINFO (1806+), SCOPUS, Web of Science, and Cochrane Library. The specific search strategy for the MEDLINE is described in Supplementary Appendix 1.

Two independent researchers (L.Y., L.R.L.) were involved in the article screening and selection. After elimination of irrelevant articles by screening of the title and abstract, the remaining articles were reviewed in full text to determine their eligibility. The reference lists of the eligible articles were also examined to find other potentially relevant articles. The Science Citation Index was used for an additional search to identify all relevant articles that referenced the eligible articles. The last search was done on 20 January 2014. The agreement between the two independent researchers in article selection was evaluated by Kappa statistic.

2.2. Selection Criteria

The inclusion criteria were: the study utilized dual-task balance performance as an outcome measurement; a clear description of the methods used to assess dual-task performance was provided; psychometric properties of a specific dual-task test were evaluated and reported; the study sample involved older adults; published in English. The exclusion criteria were: the study involved only individuals with a specific primary diagnosis (e.g., stroke); dissertation theses, review articles or conference abstracts.

2.3. Methodological Quality Assessment

Two independent researchers (L.Y., F.M.H. L.) evaluated the quality of the included articles using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist (Table 1) [21]. This tool assesses nine different domains of psychometric properties, and each domain can be rated as excellent, good, fair, or poor in accordance with the specific criteria described in the checklist. Any disagreement in ratings between these two researchers was resolved after discussion with the principal investigator (M.Y.C.P.).

2.4. Data Synthesis and Analysis

The findings were extracted by the first author (L.Y.) and confirmed with the principal investigator (M.Y.C.P.). Due to the large heterogeneity of the selected articles and different outcomes measures used, meta-analysis was not performed.

3. RESULTS

3.1. Article Selection

Twenty-six articles (23 studies) [7-12,15,22-40] were included in this review

(Figure 1). The inter-rater agreement for article selection was good (Kappa=0.77). The characteristics of study participants are summarized in Supplementary Table 1 and 2.

3.2. Dual-task Testing Protocol

Pollock et al [41] proposed the term of “walking balance” to indicate all the walking-related tasks that challenge the balance system. Therefore, we classified the primary task into two domains, namely, static balance domain, and walking balance domain. According to this classification, three studies (4 articles) evaluated dual-task static balance assessments [15,34,38,39] (Supplementary Table 1), whereas 19 studies (21 articles) involved dual-task walking balance assessments [7-11,22-33,35-37,40], one study (one article) assessed both static and walking balance [12] (Supplementary Table 2).

3.2.1 Primary Task: Static Standing Balance

A force platform [42] was used to measure static standing balance ability in three studies (four articles) [15,34,38,39]. Although the testing conditions were different, centre of pressure (COP)-related parameters (e.g., displacement, velocity) in mediolateral and anterioposterior directions were used as outcomes [15,34,38,39].

3.2.2 Primary Task: Walking Balance

Two studies investigated stepping execution [12,33]. The outcomes included step initiation, preparation and swing phase, foot-off time and foot contact time (measured by a force platform) [33]. A variety of walking tests were used as the primary task in 18 studies (20 articles) [7-10,22-29,35,36,40]. Walking speed or walking time was the outcome variable in all these studies, although variability of gait parameters was also used in some [25,30,31].

3.2.3 Secondary Task

Al-Yahya et al. [14] classified the secondary cognitive tasks into five categories, namely, mental tracking, verbal fluency, working memory, reaction time, and discrimination and decision-making. These were used in sixteen [7,8,15,22-25,29,30,34-40], six [9,27,31,32,35,36], one [10], two [12,26] and four [26,33,35,36] articles respectively. Seven [7,8,28,35-37,40] articles used a manual task as the secondary task.

3.2.4 Assessment of Dual-Task Interference

To assess dual-task interference, many studies examined whether there was a statistically significant difference in static standing or walking balance performance between single-task and dual-task conditions. Two studies (three articles) evaluated the psychometric properties of the dual-task cost (DTC) [35-37], which can be computed as absolute DTC (absolute DTC=dual-task performance minus single-task performance), or as relative DTC [relative DTC=(dual-task performance minus single-task performance)*100%/single-task performance] [10].

3.3. *Methodological Quality*

Out of nine domains of psychometric properties described in COSMIN, only four were reported in the selected articles, namely, reliability (16 articles) [7-10,12,15,25,27,29-31,33,34,36,39,40], measurement error (five articles) [27,29,30,34,39], hypothesis testing (two articles) [28,32], and criterion validity (15 articles) [7-9,11,12,15,22-24,26,27,35,37,38,40] (Table 1). For the reliability domain, none of the articles was rated as good or excellent. The ratings for the specific criteria in each domain are shown in Supplementary Table 3 to 6.

3.4. *Psychometric Properties: Dual-task Static Balance Assessments*

3.4.1 Standing with Mental Tracking (4 articles)

In general, dual-task static balance assessments using COP-related parameters demonstrated moderate to excellent reliability (mediolateral direction: ICC=0.65-0.98 [15,34]; anterioposterior direction: ICC=0.52-0.74) [39] (Table 2).

COP-related measures obtained while standing on the moveable force platform were moderately correlated with clinical measures of balance, gait, and activity levels, (>0.47) (i.e., convergent validity) [15]. In terms of predictive validity, Swanenburg et al. [38] found that none of the COP-related parameters under the dual-task condition was useful in predicting falls. Condron and Hill et al. [15], in contrast, reported that balance performance while standing on a dynamic platform with forward-backward tilting under dual-task condition had a high sensitivity and specificity (both 0.8) in discriminating a healthy older adult group from a group with mild increase in fall risk.

3.4.2 Standing with Reaction Time (1 article)

The reaction time measured in a push-button task in response to a visual stimulus during quiet standing could not discriminate fallers from non-fallers [12].

3.4.3 Standing with Reaction Time and Mental Tracking (1 article)

The reaction time in the above task while counting backward showed good reliability (ICC=0.75), but was not associated with falls [12].

3.5. Psychometric Properties: Dual-task Dynamic Balance Assessments

3.5.1. Stepping with Discrimination and Decision Making (1 article)

The COP-related parameters showed moderate to good inter- and intra-rater reliability (ICC=0.62-0.85) [33] (Table 2).

3.5.2 Stepping with Reaction Time (1 article)

The reaction time in the push-button task while stepping in place showed excellent reliability (ICC=0.90) and was significantly associated with falls (OR=3.16) [12].

3.5.3 Stepping with Mental Tracking and Reaction Time (1 article)

The reaction time in the above task while counting backward showed moderate reliability (ICC=0.66) but was not useful in identifying fallers [12].

3.5.4 Walking with Mental Tracking (12 articles)

The gait spatiotemporal parameters (walking speed, step length, etc.) had excellent reliability (ICC=0.85-0.99) [7,8,29,30,36]. In contrast, if the variability of gait parameters were used as the outcome, the results were less consistent, with tremendous difference in the reliability values reported (ICC=$0.11_{\text{(step time)}}$ to $0.88_{\text{(step length)}}$) [25,30]. The reliability of the parameters derived from the cognitive task (fluency, accuracy rate, etc.), and the DTC for the walking and cognitive tasks was only poor to moderate (ICC=0.19-0.58) [36].

The time taken to complete the TUG while engaging in mental tracking (TUGcog) was moderately correlated to the Berg Balance Scale (i.e., convergent validity) [7]. The increase in the number of enumerated figure while walking, and change in mean step-width of greater than 3.6mm, and time taken to complete TUGcog was significantly associated with falls [8,24,37]. Other studies, however, found that walking speed [11,22] and TUGcog were not significant in predicting falls [22]. TUGcog was also not useful in differentiating pre-frail individuals from non-frail ones [40].

3.5.5 Walking with Verbal Fluency (6 articles)

Similar to the mental tracking task, the most reliable parameters were derived from the walking task (walking speed, stride length, etc.), all indicating good to excellent reliability (ICC=0.78-0.92) [27,31,36,40]. The reliability of the verbal fluency task was also good (ICC>0.78) [36]. The reliability of the variability measure of the mobility task (ICC=0.226) [31], DTC for the mobility (ICC=0.59-0.67) and cognitive task (ICC<0.35) was only poor to fair [36].

Two studies assessed convergent validity and found that performance in the walking task was moderately correlated with other measures of functional performance [27,32]. Both the WWT-simple and WWT-complex had low sensitivity, but relatively high specificity in discriminating fallers [9].

3.5.6 Walking with Working Memory (1article)

The reliability for walking time was good to excellent (inter-rater: ICC>0.97; intra-rater: ICC>0.79) [10]. The number of steps off the path was also counted and good reliability was reported (Kappa=76%). Inter-rater reliability for digit span accuracy test was excellent (ICC>0.97) [10].

3.5.7 Walking with Reaction Time (1 article)

Neither walking time nor reaction time was a significant predictor of falls [26].

3.5.8 Walking with Discrimination and Decision-Making (2 articles)

The intra-rater reliability of the walking time was good (ICC=0.85) while that of the visual-spatial decision time (ICC=0.39), and the DTC for these two tasks was low (ICC<0.37) [36]. Neither walking time nor visual-spatial decision time under dual-task condition was a significant predictor of falls [26].

3.5.9 Walking with Manual Task (7 articles)

The intra-rater reliability of the time taken to complete TUG combined with a manual task (TUGman) was excellent (ICC>0.88) [7,8,36,40], but that of the DTC was only fair (ICC=0.50) [36]. The time taken to complete the TUGman was moderately correlated to the Berg Balance Scale (i.e., convergent validity) [7], and was useful in differentiating pre-frail individuals from non-frail ones [40], and fallers from non-fallers [8]. Increase or decrease in mean step-width ⩾±3.7mm in dual-task condition was associated with decreased fall risk (OR=0.2) [37]. Walking speed under dual-task condition was moderately correlated with self-reported anxiety level (r=0.4-0.5) [28].

3.5.10 Walking with Discrimination and Decision Making and Manual Task (2 articles)

The walking time had good reliability and (ICC=0.88), and significant association with falls (OR=1.12) [35,36].

4 Discussion

4.1 Reliability Analysis

4.1.1 Dual-task Static Balance Assessments

Despite different testing protocols used, COP-related parameters in the mediolateral direction tended to be more reliable (ICC=0.65-0.98) than in the anteroposterior direction (ICC=0.52-0.74) [15,34,39]. The reason underlying this phenomenon is not entirely clear but may be related to the difference in excursion of movements between the anteroposterior and mediolateral directions. In quiet standing, the ankle strategy is most likely used to maintain body equilibrium [43]. The available range of motion for the ankle complex is much greater in the anterioposterior direction (dorsiflexion/plantarflexion) than in the mediolateral direction (inversion/eversion), thus contributing to more variability in the former.

4.1.2 Dual-task Walking Balance Assessments

Generally, the reliability was high when the gait spatiotemporal parameters under dual-task conditions were used as outcomes, regardless of the type and complexity of the secondary task [10,25,27,29-31,36]. In contrast, the reliability was only fair to poor when the variability of gait parameters was used as outcomes [25,30,31]. As revealed by Hollman et al. [31], measuring variability in stride velocity would require data that are collected from approximately 220 strides to achieve a magnitude of reliability equal to 0.90. In the three studies that investigated the reliability of variability of gait parameters [25,30,31], the walking distance used was only 10m [25,31] or 18m [30], in which the number of strides involved would have been much less than 220, and thus may not provide a good indication of variability of gait, resulting in relatively low reliability. In these three studies, participants were all community-ambulating older adults, which may lead to low heterogeneity within the sample and thus lower reliability.

The reliability of DTC for the gait parameters was poor to fair, with ICC ranging from 0.37 to 0.67 for primary tasks [36], and ≤0.32 for secondary tasks [36]. It may be attributable to the inflation in the systematic error when combining two variables to calculate the difference in performance between single-task and dual-task conditions [36]. Another cause for the low reliability of DTC may be the change in posture strategy across trials. Since none of the selected studies specified the prioritization of tasks, the change in prioritization of the cognitive task (posture-second strategy) versus the balance task (posture-first strategy) may lead to different performance [44,45], which became magnified when DTC was computed.

Few studies examined the psychometric properties of the outcomes derived from the cognitive task. Generally, the reliability of the cognitive tasks was lower than the balance tasks, in both single-task and dual-task conditions. Similar phenomenon was observed when such tasks were tested in children and older adults with Alzheimer's disease [36]. One possible explanation could be that attention is a multifaceted cognitive construct, including sustained attention, shifting attention, and dividing attention [1], which may make the cognitive task performance inherently unstable. There is some evidence that the choice of the cognitive task may also influence the reliability, with the verbal fluency task (ICC>0.78) [36] and digit span accuracy (ICC>0.97) [10] showing higher reliability than others. These cognitive tasks may therefore be preferred options when assessing dual-task balance performance in older adults.

4.2 Validity Analysis

4.2.1 Dual-task Static Balance Assessments

The COP-related parameters [18] and reaction time measurements [12] in single or dual-task conditions were not significant in predicting falls. As only quiet standing was assessed, it may not pose sufficient challenge to the postural control capacity to differentiate fallers from non-fallers [38]. Indeed, most falls among older adults occur during dynamic activities (e.g., walking, reaching/leaning), rather than static standing [46].

4.2.2 Dual-task Walking Balance Assessments

The convergent validity of the dual-task walking balance assessment was good, regardless of whether the TUG or other walking tests were used as the mobility task [7,27,28,32].

Out of ten studies [8,9,11,12,22-24,26,35,37] that investigated the relationship between dual-task walking balance performance and falls, six studies [8,9,12,23,24,35,37] reported significant results. In general, poorer performance in dual-task walking (slower walking speed) was correlated with increased falls. In six studies, the sensitivity and specificity values derived from the walking balance tests under single-task condition were also reported [8,9,22-24,35]. However, it remains unclear as to whether the dual-task assessment is more useful than single-task assessment in fall prediction. Two studies showed that single-task performance and dual-task performance had similar predictive values [8,35]. Three studies demonstrated that outcomes from the dual-task performance are better in predicting falls than single-task performance [9,23,24], but not others [8,22,35].

Nordin et al. [37] demonstrated that the absolute magnitude of DTC, rather than the direction of change in performance, was an important predictor of falls. In their study, DTC in step-width greater than ±3.6mm while performing serial 3 subtractions simultaneously was indicative of increased fall risk whereas DTC in the same variable greater than ±3.7mm while carrying a cup was associated with reduced fall risk. Thus, the former was considered as an inadequate strategy (risk factor) while the latter a protective strategy. Their results suggested that different cognitive load imposed may affect the prioritization during walking balance performance.

Tang et al. [40] demonstrated that TUGman was more sensitive than the TUG test alone in identifying pre-frail individuals. With the addition of a manual task, TUGman demands a higher level of dynamic balance control and challenges sustained attention while maintaining the grip force [47]. As muscle weakness was one of the key signs of frailty [48], it may partially explain why TUGman was more useful in identifying fallers.

4.3 Limitations of the Studies Reviewed

The methodological quality was rated as “fair” and “poor” in the “Reliability” and “measurement Errors” domains, mainly due to the small sample size used (i.e.⩽30) [8,10,15,27,29-31,33,34,39]. None of the studies examined responsiveness of dual-task assessments, which should be explored in future studies.

4.4 Limitations of this Systematic Review

In most of the studies, the subjects were living independently and had normal cognition. The findings were not generalizable to older people with different characteristics. There is increasing evidence that cognitive impairment has an important role in dual-task performance [11,13,25]. It is thus possible that the psychometric properties of dual-task assessment tools may differ depending on the cognitive status. Hence, there is a need for better understanding of dual-task measures across varying cognitive conditions (e.g., Alzheimer's disease, Parkinson's disease, etc.). Lastly, no meta-analysis was done to estimate the overall OR value for fall prediction, due to the vast differences in the assessment protocol, definition of fallers, and study design.

5 Conclusion

In assessing static standing balance under dual-task condition, the COP-related parameters in the mediolateral direction showed good reliability, but were not useful in predicting falls. For assessment of dual-task walking balance, walking speed or walking time also showed good reliability but the results on fall prediction were mixed. Outcomes derived from the secondary task and DTC generally showed poorer reliability. The findings in this review should be useful in guiding the selection and design of dual-task balance outcome measures in future research.

Contributors

All authors contributed to the manuscript. Dr. Pang, Mr. Yang and Dr. He provided concept/idea/research design. Dr. Pang and Mr. Yang provided writing. Dr. Pang, Mr. Yang, and Mr. Liao provided data collection. Mr. Yang, Mr. Lam, and Dr. Pang provided methodological quality assessment for the included articles. Dr. Pang, Mr. Yang provided data analysis. Dr. Pang provided project management. All authors provided consultation (including review of manuscript before submission).

Competing interest

All authors declare no conflict of interest.

Funding

This work was supported by a PhD research studentship provided by the Institute for Disaster Management and Reconstruction, Sichuan University – Hong Kong Polytechnic University [RTHV] [L. Y.], and by the General Research Fund provided by the Research Grants Council [no. PolyU 5245/11E] [L-R. L., M.Y.C.P.]. F. M. H. L. was supported by a PhD full-time studentship from the Hong Kong Polytechnic University [RT5F]. The funding agency played no role in the design, execution, analysis and interpretation of data, or writing of the study.

Provenance and peer review

Not commissioned and externally peer-reviewed.

REFERENCES

1. McCulloch K. Attention and dual-task conditions: physical therapy implications for individuals with acquired brain Injury. J Neurol Phys Ther 2007;31:104-18.
2. Cesari M, Onder G, Russo A, Zamboni V, Barillaro C, Ferrucci L, et al. Comorbidity and physical function: results from the aging and longevity study in the Sirente geographic area (ilSIRENTE study). Gerontology 2006;52:24-32.
3. Yeung P, Towers A, La Grow S, Philipp M, Alpass F, Stephens C. Mobility, satisfaction with functional capacity and perceived quality of life (PQOL) in older persons with self-reported visual impairment: the pathway between ability to get around and PQOL. Disabil Rehabil 2014. doi:10.3109/09638288.2014.910559.
4. Patla AE, Shumway-Cook A. Dimensions of mobility: defining the complexity and difficulty associated with communitymobility. J Aging Phys Act 1999;7:7-19.
5. De Hoon E, Allum J, Carpenter M. Quantitative assessment of the stops walking while talking test in the elderly. Arch Phys Med Rehabil 2003;84:838-42.
6. Lundin-Olsson L, Nyberg L, Gustafson Y. “Stops walking when talking” as a predictor of falls in elderly people. Lancet 1997;349:617.
7. Hofheinz M and Schusterschitz C. Dual task interference in estimating the risk of falls and measuring change: a comparative, psychometric study of four measurements. Clin Rehabil 2010;24:831-42.
8. Shumway-Cook A, Brauer S, Woollacott M. Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. Phys Ther 2000;80:896-903.

9. Verghese J, Buschke H, Viola L. Validity of divided attention tasks in predicting falls in older individuals: A preliminary study. J Am Geriatr Soc 2002;50:1572-6.

10. McCulloch KL, Mercer V, Giuliani C, Marshall S. Development of a clinical measure of dual-task performance in walking: reliability and preliminary validity of the Walking and Remembering Test. J Geriatr Phys Ther 2009;32:2-9.

11. Casas-Herrero A, Cadore EL, Zambom-Ferraresi F, Idoate F, Millor N, Martinez-Ramirez A, et al. Functional capacity, muscle fat infiltration, power output, and cognitive impairment in institutionalized frail oldest old. Rejuvenation Res 2013;16:396-403.

12. Makizako H, Furuna T, Shimada H, Ihira H, Kimura M, Uchiyama E, et al. Association between a history of falls and the ability to multi-task in community-dwelling older people. Aging Clin Exp Res 2010;22:427-32.

13. Montero-Odasso M, Muir SW, Speechley M. Dual-task complexity affects gait in people with mild cognitive impairment: the interplay between gait variability, dual tasking, and risk of falls. Arch Phys Med Rehabil 2012;93:293-9.

14. Al-Yahya E, Dawes H, Smith L. Cognitive motor interference while walking: A systematic review and meta-analysis. Neurosci Biobehav Rev 2011;35:715-28.

15. Condron JE and Hill KD. Reliability and validity of a dual-task force platform assessment of balance performance: effect of age, balance impairment, and cognitive task. J Am Geriatr Soc 2002;50:157-62.

16. Kelly VE, Schrager MA. Age-associated effects of a concurrent cognitive task on gait speed and stability during narrow-base walking. J Gerontol A Biol Sci Med Sci 2008;63:1329-34.

17. Beauchet O, Annweiler C, Dubost V. Stops walking when talking: a predictor of falls in older adults? Eur J Neurol 2009;16:786-95.

18. Chu Y, Tang P, Peng Y, Chen H. Meta-analysis of type and complexity of a

secondary task during walking on the prediction of elderly falls. Geriatr Gerontol Int 2013;13:289-97.

19. Zijlstra A, Ufkes T, Skelton DA, Lundin-Olsson L. Do dual tasks have an added value over single tasks for balance assessment in fall prevention programs? A mini-review. Gerontology 2008;54:40-9.

20. Ashford S, Slade M, Malaprade F, Turner-Stokes L. Evaluation of functional outcome measures for the hemiparetic upper limb: a systematic review. J Rehabil Med 2008;40:787-95.

21. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012;21:651-7.

22. Beauchet O, Allali G, Annweiler C, Berrut G, Maarouf N, Herrmann FR, et al. Does change in gait while counting backward predict the occurrence of a first fall in older adults? Gerontology 2008;54:217-23.

23. Beauchet O, Annweiler C, Allali G, Berrut G, Herrmann F, Dubost V. Recurrent falls and dual task-related decrease in walking speed: is there a relationship? J Am Geriatr Soc 2008;56:1265-69.

24. Beauchet O, Dubost V. ‘Faster counting while walking’ as a predictor of falls in older adults. Age Ageing 2007;36:418-23.

25. Beauchet O, Freiberger E, Annweiler C, Kressig RW, Herrmann FR, 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.

26. Faulkner KA, Redfern MS, Cauley JA, Landsittel DF, Studenski SA, Rosano C, et al. Multitasking: Association between poorer performance and a history of recurrent falls. J Am Geriatr Soc 2007;55:570-6.

27. Gimmon Y, Jacob G, Lenoble-Hoskovec C, Büla C, Melzer I. Relative and absolute reliability of the clinical version of the Narrow Path Walking Test (NPWT) under single and dual task conditions. Arch Gerontol Geriatr

2013;57:92-9.

28. Hadjistavropoulos T, Carleton RN, Delbaere K, et al. The relationship of fear of falling and balance confidence with balance and dual tasking performance. Psychol Aging 2012;27:1-13.

29. Hars M, Herrmann FR, Trombetti A. Reliability and minimal detectable change of gait variables in community-dwelling and hospitalized older fallers. Gait Posture 2013;38:1010-4.

30. Hartmann A, Murer K, De Bie RA, De Bruin ED. Reproducibility of spatio-temporal gait parameters under different conditions in older adults using a trunk tri-axial accelerometer system. Gait Posture 2009;30:351-5.

31. Hollman JH, Childs KB, McNeil ML, Mueller AC, Quilter CM, Youdas JW. Number of strides required for reliable measurements of pace, rhythm and variability parameters of gait during normal and dual task walking in older individuals. Gait Posture 2010;32:23-8.

32. Liu-Ambrose T, Katarynych LA, Ashe MC, Nagamatsu LS, Hsu CL. Dual-task gait performance among community-dwelling senior women: the role of balance confidence and executive functions. J Gerontol A Biol Sci Med Sci 2009;64:975-82.

33. Melzer I, Shtilman I, Rosenblatt N, Oddsson LI. Reliability of voluntary step execution behavior under single and dual task conditions. J Neuroeng Rehabil 2007;4:16.

34. Moghadam M, Ashayeri H, Salavati M, Sarafzadeh J, Taghipoor KD, Saeedi A, et al. Reliability of center of pressure measures of postural stability in healthy older adults: effects of postural task difficulty and cognitive load. Gait Posture 2011;33:651-5.

35. Muhaidat J, Kerr A, Evans JJ, Pilling M, Skelton DA. Validity of simple gait-related dual-task tests in predicting falls in community-dwelling older adults. Arch Phys Med Rehabil 2014;95:58-64.

36. Muhaidat J, Kerr A, Evans JJ, Skelton DA. The test-retest reliability of gait-related dual task performance in community-dwelling fallers and

non-fallers. Gait Posture 2013;38:43-50.

37. Nordin E, Moe-Nilssen R, Ramnemark A, Lundin-Olsson L. Changes in step-width during dual-task walking predicts falls. Gait Posture 2010;32:92-7.

38. Swanenburg J, De Bruin E, Uebelhart D, Mulder T. Falls prediction in elderly people: A 1-year prospective study. Gait Posture 2010;31:317-21.

39. Swanenburg J, De Bruin ED, Favero K, Uebelhart D, Mulder T. The reliability of postural balance measures in single and dual tasking in elderly fallers and non-fallers. BMC Musculoskelet Disord. 2008;9:162.

40. Tang PF, Yang HJ, Peng YC, Chen HY. Motor dual-task Timed Up & Go test better identifies prefrailty individuals than single-task Timed Up & Go test. Geriatr Gerontol Int 2014. doi: 10.1111/ggi.12258.

41. Pollock C, Eng J, Garland S. Clinical measurement of walking balance in people post stroke: a systematic review. Clin Rehabil 2011;25:693-708.

42. Lin C-C, Whitney SL. Quantification of Static and Dynamic Balance While Maintaining and Changing Body Position. Top Geriatr Rehabil 2012;28:17-26.

43. Shumway-Cook A, Woollacott M, Motor control: translating research into clinical practice. 3rd ed. Baltimore, MA: Lippincott Williams & Wilkins; 2007.

44. Shumway-Cook A, Woollacott M, Kerns K, Baldwin M. The effects of two types of cognitive tasks on postural stability in older adults with and without a history of falls. J Gerontol A Biol Sci Med Sci 1997;52:M232-40.

45. Bloem BR, Grimbergen YA, van Dijk JG, Munneke M. The "posture second" strategy: a review of wrong priorities in Parkinson's disease. J Neurol Sci 2006;248:196-204.

46. Nachreiner NM, Findorff MJ, Wyman JF, McCarthy TC. Circumstances and consequences of falls in community-dwelling older women. J Womens Health (Larchmt) 2007;16:1437-46.

47. Lundin-Olsson L, Nyberg L, Gustafson Y. Attention, frailty, and falls: the effect of a manual task on basic mobility. J Am Geriatr Soc 1998;46:758-61.

48. Evenhuis HM, Hermans H, Hilgenkamp TI, Bastiaanse LP, Echteld MA.

Frailty and disability in older adults with intellectual disabilities: results from the healthy ageing and intellectual disability study. J Am Geriatr Soc 2012;60:934-8.

FIGURE LEGENDS

Figure 1. Flow diagram.

The diagram illustrates the flow of information through the different phases of the systematic review. A total of 26 articles (23 studies) were included in this review.

Table 1. Evaluation of methodological quality using the COSMIN Checklist

		Total number of articles	Number of articles (Percentage) [reference number]			
Box	**Domain**		**Poor**	**Fair**	**Good**	**Excellent**
A	Internal consistency					
B	Reliability	16	10 (62%) [7,9,10,12,15,30,31,33,34,40]	6 (38%) [8,25,27,29,36,39]		
C	Measurement error	5	2 (40%) [30,34]	3 (60%) [27,29,39]		
D	Content validity					
E	Structural validity					
F	Hypotheses testing	2		2 (100%) [28,32]		
G	Cross-cultural validity					
H	Criterion validity	15	1 (7%) [11]	6 (40%) [8,9,12,15,26,27]	3 (20%) [35,38,40]	5 (33%) [7,22-24,37]
I	Responsiveness					

Table 2. Psychometric Properties of Dual-task Balance Assessments: Summary

Dual-Task Condition (Number of articles, N)	**Reliability**		**Validity**	
	Intraclass Correlation Coefficient (ICC)	**Other Measures of Reliability**	**Concurrent**	**Predictive**
Dual-Task Static Balance Assessment				
Static Standing balance + Mental Tracking (N=4)	COP-related parameters (displacement, velocity): Intra-rater: ML direction: ICC>0.65*[15] ICC=0.70~0.98*[34,39] AP direction: ICC=0.54~0.63*[39] Inter-rater: ML direction: ICC=0.71~0.86*[39] AP direction: ICC=0.46~0.72*[39]	COP-related parameters (displacement, SD of amplitude, SD of velocity, phase plane): SEM: Parameters in the ML direction < in the AP direction, [34] MDC: Max-ML= 0.37cm [39], Max-AP=0.83cm [39], AoE=1.48~3.75cm^2 [39] Parameters in the ML direction < in the AP direction [34]	COP-related parameters: Correlation with balance (step test and timed up and go), activity levels (HAP), and some gait measures (gait velocity and stride length): Pearson's r >0.47*[15]	COP-related parameters not significant in predicting falls [38]
Static Standing Balance + Reaction time				Reaction time: no association with falls (p>0.05) [12]

Static Standing Balance + Mental Tracking + Reaction Time (N=1)	Reaction time performance: Intra-rater: ICC=0.75*[12]			Delay ratio of reaction time response to predict falls: OR(95%CI)=1.21 (0.78, 1.85) (NS) [12]
Dual-Task Walking Balance Assessment				
Stepping + Discrimination and Decision-making (N=1)	Five temporal parameters: step initiation, preparation and swing phases, foot-off time and foot contact time): Inter-rater: ICC=0.70~0.83*[33] Intra-rater: ICC=0.62~0.85*[33]	Internal consistency, Cronbach's alpha: Intra-rater: highest (foot-off time)=0.94 [33], lowest (swing phase duration)=0.59~0.64 [33] Inter-rater: highest (foot-contact and foot-off times) = 0.85~0.91[33] lowest (swing phase duration) = 0.48~0.78 [33]		
Stepping + Reaction Time (N=1)	Reaction time performance: Intra-rater: ICC=0.90* [12]			Delay ratio of reaction time response to predict falls: OR(95%CI)=3.16 (1.06, 9.45) [12]
Stepping + Mental Tracking + Reaction Time (N=1)	Reaction time performance: Intra-rater: ICC=0.66* [12]			Delay ratio of reaction time response to predict falls: OR(95% CI)=1.20 (0.82, 1.74) (NS) [12]

TUG /Walking + Mental Tracking (N=12)	TUGcog completion time: Intra-rater: ICC=0.94*[7] ICC =0.98[40] Inter-rater: ICC=0.99*[8] Walking speed, cadence, step duration, step length: intra-rater: ICC=0.86-0.95*[30] Inter-rater: ICC=0.90-0.97*[30] Stride time (mean value): Intra-rater: ICC=0.86*[25] Walking speed: Intra-rater: ICC=0.85*[29] Walking time: Intra-rater: ICC=0.89*[36] Cognitive task: fluency: Intra-rater:	Walking speed, cadence, step duration and step length**:** RLOAs=3%-10%[30]; Walking velocity: SEM=5.6cm/s [29], MDC$_{95}$=15.5cm/s [29], SEM% < 6.5% [29], MDC$_{95\%}$ <18.1% [29]	Correlation with TUGcog completion time: BBS: r=-0.66*[7] grip strength: r = −0.229(NS) [40]	TUGcog completion time to predict falls: Sensitivity=80%*[8] Specificity=93%*[8] Cutoff=15s [8] TUGcog completion time to predict prefrail individuals AUC (95%CI)= 0.60 (0.46, 0.74) Sensitivity=29% Specificity=93% Age-adjusted OR (95% CI)= 2.8 (0.5, 15.4) (NS) [40] Association between TUGcog completion time and incidence of falls: Spearman’s rho=0.39(NS) [11] Enumerated figures to predict falls: OR=53.3*[24] Specificity=90.0%*[24] Sensitivity=86.5%*[24]

	ICC=0.58*[36] accuracy rate: Intra-rater: ICC= 0.51*[36]	PPV=85.2% [24] NPV=90.2% [24] Walking speed to predict recurrent falls: Crude OR=0.60 (0.41, 0.85) *[23]
	DTC: Walking time: Intra-rater: ICC= 0.53*[36] Cognitive task: Fluency: Intra-rater: ICC=0.14(NS) [36] Accuracy rate: Intra-rater: ICC=0.19*[36]	Association between walking speed and incidence of falls: Spearman’s rho=0.26 (NS) [11] DTC in mean step-width to predict falls: OR=2.5* (1.1, 5.7) [37]
	Gait Variability: Intra-rater: Step duration: $ICC_{(\text{gym floor})}$=0.46*[30] $ICC_{(\text{foam})}$=0.73*[30] Step length: $ICC_{(\text{gym floor})}$=0.37*[30] $ICC_{(\text{foam})}$=0.30*[30] Stride time:	OR in walking speed not significant to predict first falls [22]

	ICC=0.11*[25] Inter-rater: Step duration: ICC$_{\text{(gym floor)}}$=0.72*[30] ICC$_{\text{(foam)}}$=0.80*[30] Step length: ICC$_{\text{(gym floor)}}$=0.71*[30] ICC$_{\text{(foam)}}$=0.88*[30]			
Walking + Verbal Fluency (N=6)	Walking speed: Intra-rater: ICC =0.89-0.93*[27,31]; Walking time: Intra-rater: ICC$_{\text{(naming animals)}}$=0.89* [36] ICC$_{\text{(alternative letters)}}$=0.92*[36] Number of step errors: Intra-rater: ICC = 0.92*[27] Cognitive task: Fluency: Intra-rater: ICC$_{\text{(alternative letter)}}$=0.78*[36] ICC$_{\text{(naming animals)}}$=0.79*[36]	Stride velocity: SEM= 0.03m/s [27], SRD=0.78m/s [27] Number of step errors: SRD=3.4 [27]	Correlation with gait velocity (Pearson's r): POMA, r = 0.38*[27] SPPB, r = 0.56*[27]; walking time with ABC: WWT-simple, r= -0.55*[32] WWT-complex, r= -0.54*[32]	Walking time to predict falls: WWT-simple: Sensitivity=46.1%*[9] Specificity=89.4%*[9] PPV=54.5% [9] OR=7.02* [9] WWT-complex: Sensitivity=38.5%*[9] Specificity=95.6%*[9] PPV=71.4% [9] OR=13.7*[9]

Accuracy rate:
Intra-rater:
ICC$_{\text{(alternative letter)}}$=0.37*[36]
ICC$_{\text{(naming animals)}}$=0.53*[36]

DTC (relative):
Walking time:
Intra-rater:
ICC$_{\text{(naming animals)}}$=0.59*[36]
ICC$_{\text{(alternative letter)}}$=0.67*[36]
Cognitive task:
Fluency:
Intra-rater:
ICC$_{\text{(alternative letter)}}$=-0.04(NS) [36]
ICC$_{\text{(naming animals)}}$=0.32*[36]
Accuracy rate:
Intra-rater:
ICC$_{\text{(alternative letter)}}$=0.19(NS) [36]
ICC$_{\text{(naming animals)}}$=0.33*[36]

Variability:
Stride velocity:
Intra-rater:

	ICC=0.226*[31]	
Walking + Working Memory (N=1)	Walking time: Inter-rater: ICC > 0.97*[10] Intra-rater: ICC > 0.79*[10] Steps off the path: Kappa=76%*[10] Digit span accuracy: Inter-rater: ICC > 0.97*[10]	
Walking+ Reaction Time (N=1)		Walking time (Turn walk) to predict falls: OR=1.23 (NS) [26] Reaction time to predict falls: OR=0.84 (NS) [26]

Walking + Discrimination and decision-making (N=2)	Walking time: Intra-rater: ICC=0.85*[36] Cognitive task: Fluency: Intra-rater: ICC=0.39*[36] Response accuracy: Intra-rater: ICC=0.24(NS) [36] DTC: Walking time: Intra-rater: ICC=0.37(NS) [36] Cognitive task performance: Fluency: Intra-rater: ICC= 0.11(NS) [36] Accuracy rate: Intra-rater: ICC=-0.04(NS) [36]		Walking time (Turn walk) to predict falls: OR=1.21 (NS) [26] Reaction time to predict falls: OR= 0.91 (NS) [26]
TUG/Walking+ Manual (N=7)	TUGman completion time: Intra-rater: ICC>0.88*[7,36]	Correlation with TUGman completion time: BBS: r=-0.72*[7]	TUGman completion time to predict falls: Sensitivity=80%*[8]

	ICC=0.99[40] Inter-rater: ICC=0.99*[8] DTC: TUGman Completion time: Intra-rater: ICC=0.50*[36]		grip strength: r = −0.283*[40] Correlation between self-reported anxiety level and walking speed: floor/tray: r= -0.44*[28] platform/tray: r= -0.45*[28]	Specificity=93%*[8] Cutoff=14.5s [8] TUGman completion time to predict prefrail individuals AUC (95%CI)= 0.73 (0.60–0.86) Sensitivity=83% Specificity=64% Age-adjusted OR (95% CI)= 7.2 (1.9, 27.6) [40] DTC of TUGman to predict falls: Completion time: OR(95% CI)=1.23* (1.10, 1.41) [35] Absolute DTC in completion time: OR(95% CI)=0.61* (0.43, 0.81)[35] DTC in mean step-width to predict falls: OR(95% CI)=0.2* (0.1, 0.5) [37]

Walking + Discrimination and decision-making + Manual (N=2)	Walking time: Intra-rater: ICC=0.88* [36] Cognitive task: Fluency: Intra-rater: ICC=0.12 [36] Response accuracy: Intra-rater: ICC=0.51* [36] DTC: Walking time: Intra-rater: ICC=0.64 [36] Cognitive task performance: Fluency: Intra-rater: ICC= 0.34* [36] Accuracy rate: Intra-rater: ICC=-0.30* [36]		Walking time to predict falls: OR(95% CI)= 1.12* (1.03, 1.24) [35]

* Statistically significant; NS: non-significant

ABC: Activity-Specific Balance Confidence Scale; AoE: area of the 95% confidence ellipse; AP: anterioposterior; BBS: Berg Balance Scale; COP: center of pressure; DTC: dual task cost; HAP: the Human Activity Profile; MCID: minimal clinical important difference; ML: mediolateral; NPV: negative predictive value; NS: not significant; OR: odds ratio; POMA: Performance Oriented Mobility Assessment; PPV: positive predictive value; RLOA: ratio limits of agreement; SDD: smallest detectable difference; SEM: standard error of measurement; SPPB: Short Physical Performance Battery; SRD: smallest real difference; TUGcog: timed up-and-go with cognitive task; TUGman: timed up-and-go with manual task; WWT test: Walking When Talking Test.