Title: Psychometric Properties of Dual-Task Balance and Walking Assessments for Individuals with Neurological Conditions: A Systematic Review

Authors:

Lei Yang [a,b], Freddy Man Hin Lam [b], Lin Rong Liao [b], Mei Zhen Huang [b], Cheng Qi He [c], Marco Yiu Chung 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, China

***Corresponding Author:**

Marco Y. C. Pang, Department of Rehabilitation Sciences, Hong Kong Polytechnic University, Hong Kong, China. Tel: 852-2766-7156 Fax: 852-2330-8656; Email: Marco.Pang@polyu.edu.hk

Total Word Count: 4207

Running Head: Dual-Task Assessment

ABSTRACT

Background: The ability of performing a balance or walking task in conjunction with a secondary cognitive or motor task, referred to as dual-task (DT) ability, is essential in daily living. While there is some evidence that DT performance is impaired in individuals with neurological conditions, using reliable and valid tools to measure DT performance is essential. This systematic review aimed to evaluate the psychometric properties of DT balance and walking assessments in individuals with different neurological conditions.

Methods: A systematic literature search was conducted using PubMed, CINAHL, MEDLINE, PsycINFO, SCOPUS, Web of Science, and Cochrane Library (last search done in April 2016). The methodological quality was rated using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist.

Results: Twenty-three articles involving individuals with stroke, Parkinson's disease, mild cognitive impairment, dementia, Alzheimer's disease, and multiple sclerosis were included. Outcomes derived from the walking tasks under DT condition generally demonstrated good reliability (correlation coefficient ≥0.75) across different neurological disorders, but their usefulness in distinguishing fallers from non-fallers was inconclusive. The reliability of outcomes derived from the cognitive/motor tasks and from the dual-task effect (DTE) (i.e., DT performance minus single-task performance) seemed to be lower but was understudied. The reliability of static or dynamic sitting/standing balance outcomes in DT condition was not assessed in any of the selected studies.

Conclusions: The reliability of the outcomes derived from walking tasks was good. The psychometric properties of other DT outcomes need to be further investigated.

Keywords: psychometrics; dual-task; balance; gait; neurological; systematic review

(Word Count: 248)

1. Introduction

The ability to walk while performing a secondary cognitive or motor task (i.e., dual-tasking) is essential in daily living. Crossing the street while carrying out a conversation and walking while holding a glass of water are just some examples of frequently encountered dual-task (DT) scenarios in everyday life.

Research evidence showed that DT performance in balance and walking was decreased after the onset of neurological disease. For example, Hyndman et al. showed that when performing a walking task in conjunction with a cognitive recall task (remembering a seven item-shopping list), significantly increased walking time and deterioration in cognitive performance was found in the stroke group, but not in the control group [1]. In individuals with Parkinson's disease (PD), when asked to walk while performing a digit span task, the walking speed was reduced by 10% [2].

The issue of DT balance and mobility deficits in individuals with neurological disorders has attracted much attention in research in the past few years, as reflected by a substantial increase in the number of publications in this area. However, to ensure accurate measurement of DT balance and walking function, studying the psychometric properties (i.e., reliability, validity, responsiveness, etc.) of the DT balance and walking assessment tools are important. The psychometric properties of DT balance and walking assessments in older adults were examined in a recent systematic review [3]. Good reliability was found for the balance and gait parameters under DT condition, while poor-moderate reliability was found for the cognitive parameters and outcomes derived from the dual-task effect (DTE) (i.e., DT performance minus single-task performance). The usefulness of DT balance and walking assessment paradigms to predict falls (i.e., criterion-related validity) was examined in two systematic reviews [3, 4], but the conclusions drawn were not consistent.

The manifestations of balance and gait deficits could be very different among different neurological disorders. The degree of cognitive-motor interference may also be affected by the extent of motor and cognitive deficits [5, 6]. Therefore, the results of the systematic review of literature on older adults may not be generalizable to individuals with neurological conditions. Currently, no systematic review has evaluated the psychometric properties of existing DT balance and walking assessment tools used in individuals with neurological disorders. This systematic review was aimed to address this gap of knowledge.

2. Methods

2.1.Study Selection

Articles were eligible if: the primary task used in the DT testing protocol was a balance or walking task; the secondary attention-demanding task was a cognitive task or a motor task; the DT testing protocol was adequately described so that it could be replicated; the performance of the primary balance/walking task or that of the secondary cognitive/motor task was used as an outcome measurement; the psychometric properties of a specific DT test was evaluated; individuals with a neurological condition as the primary diagnosis were studied; and published in English. The exclusion criteria were: specific diagnosis was not provided for the study sample; conference abstracts, dissertation thesis or review articles.

2.2.Data Sources and Searches

To identify all relevant articles, a systematic literature search was conducted, involving MEDLINE (1965-9 April 2016), PubMed, SCOPUS, Web of Science, CINAHL (1982-9 April 2016), PsycINFO (1806+), and Cochrane Library electronic databases. Supplementary Appendix 1 provides an example of the search strategy.

Article screening and selection was performed by two independent researchers. The title and abstract of articles were first screened to exclude articles that did not fulfill the selection criteria. Next, the full text of the rest of articles was reviewed to decide whether these articles should be included for review. To find other potential relevant articles, the reference list of each selected article was reviewed. In addition, a forward search was conducted using the Science Citation Index to obtain other relevant articles that had cited the articles originally selected. The last forward search was done on 29 April 2016. Kappa statistics was used to assess the degree of agreement in article selection between the two independent researchers.

2.3.Methodological Quality Assessment

The Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist (Table 1) was used to assess the methodological quality of each selected article [7]. Nine different domains of psychometric properties could be assessed with this tool, in which the rating score was excellent, good, fair, or poor according to the specific criteria described in COSMIN. The rating was performed independently by two researchers. If there were discrepancies in ratings, the principal investigator was consulted to reach a consensus.

2.4.Data Synthesis and Analysis

Each article was carefully reviewed and the key findings were extracted by the first author. The results were then sent to the principal investigator for final review. For analysis of internal consistency, the interpretation of Cronbach's alpha value was as follows: <0.7=low internal consistency; 0.7-0.9=good internal consistency; >0.9=indicative of item redundancy [8]. For interpreting the reliability coefficients, [e.g., intraclass correlation coefficient (ICC), Kenadall's coefficient of concordance, Krippendorff's alpha], a value of <0.75 was regarded as poor-moderate whereas a value ≥0.75 was considered as good [8]. For Kappa values in reliability analysis, a magnitude of >0.8 was considered as excellent, 0.61-0.80 as good, 0.40-

0.60 as moderate, and <0.40 as poor to fair [8]. The interpretation of correlation coefficients (e.g., Pearson's, Spearman's) were based on the following guidelines: 0.00-0.25=little or no relationship; 0.25-0.50=fair relationship; 0.50-0.75=moderate-good relationship; >0.75=good-excellent relationship. For receiver operating characteristics (ROC) curve analysis, an area under curve (AUC) value of <0.7 represents poor discrimination; 0.7–0.8 acceptable discrimination; 0.8–0.9 excellent discrimination; 0.9–1.0 outstanding discrimination [9].

Meta-analysis was not conducted, because of the large heterogeneity of the DT balance and walking assessment tools and outcomes used. To facilitate easier comparisons of results, the findings were tabulated according to the type of neurological conditions studied, and the type of cognitive tasks (i.e., mental tracking, verbal fluency, working memory, and discrimination and decision-making) or motor tasks used [3].

3. Results

3.1.Article Selection

Twenty-three articles (23 studies) [1,2,5,6,10-28] were included in this review (Figure 1). The neurological conditions studied included stroke (8 studies) [1,5,10,12,16,17,22,28], PD (8 studies) [2,13,14,18,24-27], mild cognitive impairment (MCI), Alzheimer's disease (AD), dementia (4 studies) [6,12,19,21], and multiple sclerosis (MS) (3 studies) [15,20,23]. The inter-rater agreement for article selection was good (Kappa=0.74). The participants' characteristics and DT testing protocol for each study are summarized in Table 2.

3.2.Dual-task Testing Protocol

A walking task was used in the dual-task testing protocol for all selected studies, except that unsupported sitting balance in individuals with stroke was studied in Harley et al. [16], and that static standing balance in individuals with PD was investigated in Barbosa et al. [25]. The GAITRite system was used to measure the gait parameters in seven studies [2,11,12,18-20,23]. Different walking tests were adopted in other studies, including 20-meter walking (1 study) [27], 10-meter walking (4 studies) [5,6,15,28], 5-meter turn walking (2 studies) [1,21], and the timed up-and-go (TUG) test (4 studies) [10,22,24,28]. Stepping-in-place was also used in one study [26].

Different categories of cognitive tasks were used in the DT testing protocol, including mental tracking (9 studies) [6,11,12,15,19,22,24,27,28], verbal fluency (12 studies) [5,10,13,14,16,17,20,21,23,25,26,28], working memory (3 studies) [1, 2, 26], and discrimination and decision-making (3 studies) [2,5,18]. A secondary motor task was used in eight studies, including the task of carrying a cup water (5 studies) [5,10,18,20,28], horizontal head turns (1 study) [22], box-checking (1 study) [27], and a mobile phone task (1 study) [2]. In none of the studies were participants instructed to prioritize either the primary balance/walking task or secondary cognitive/motor task, except the study by Lee et al. [5], in which the secondary tasks were prioritized.

3.3.Assessment of Dual-Task Interference

DT interference can be assessed by comparing the performance in the single-task (ST) and DT conditions. DT interference may also be expressed as the dual-task cost (DTC) or dual-task effect (DTE), which was evaluated in six studies [2,5,14,22,23,28]. The DTE was expressed as the absolute DTE (absolute DTE=DT performance minus ST performance) [14,28], or the DTE% [DTE%=(DT performance minus ST performance)*100%/ST performance] [2,5,23,27,28]. In Tsang et al. [22], the differences between ST and DT performance were rated on an ordinal scale [items 11 and 14 of the Mini-Balance Evaluation Systems Test (Mini-BESTest); 0: normal, no noticeable change in speed (items 11 and 14); 1: Moderate interference: performs head turns with reduction in gait speed (item 11), change in speed/counting >10% (item 14), 2: Severe interference: performs head turns with imbalance (item 11), stops counting while walking or stops walking while counting (item 14)] [22].

3.4.Methodological Quality

Five out of nine domains of psychometric properties described in the COSMIN checklist were investigated in the selected articles, including internal consistency (1 study) [13], reliability (9 articles) [2,11-13,19-22,28], measurement error (5studies) [2, 12, 20, 21, 28], hypothesis testing (7 studies) [5,6,14,16,23,25,28], and criterion validity (8 studies) [1,10,15,17,18,26-28]. The COSMIN ratings for each article were summarized in Table 1. Two or more domains of psychometrics properties were assessed in some studies, and it was possible that different domains may receive different COSMIN ratings in a given article. None of the articles were rated as good or excellent in the internal consistency, reliability and measurement error domains according to the COSMIN criteria. Details of the rating of each criterion under each domain can be found in the supplementary Table 1-5 (online only).

3.5.Psychometric Properties Specific to Each Neurological Condition

3.5.1. Stroke (8 studies)

3.5.1.1.Walking

Reliability: Two studies evaluated the reliability of the outcomes derived from DT walking assessments [12,28]. With the exception of cadence (ICC=0.69), other gait parameters (velocity, step length, and stride length) yielded good intra-rater reliability (ICC=0.79-0.93) [12,28]. The reliability of the cognitive tasks [correct response rate (CRR) generated in both mental tracking and verbal fluency tasks (ICC=0.58-0.81)], and manual task (carrying a glass of water) (Kappa = 0.18–0.54) in DT condition, and

DTE and DTE% in walking time (ICC=0.11-0.80) was more varied [28]. The reliability of DTE or DTE% for the CRR was poor (ICC=0.31–0.40) [28]. DTE% was tested in item 11 and 14 of the Mini-BESTest, with moderate to good intra-rater and inter-rater reliability (Kappa=0.41-0.76) [22].

The absolute reliability of several DT walking assessments was also established by Yang et al. [28] (Table 3). The percentage minimal detectable changes at the 95% confidence level (MDC_{95}%) were generally in the 29%-45% range for the walking time outcomes. These values were substantially larger for other outcomes such as the CRR, DTE, and DTE% [28].

Validity: Several studies explored the ability of DT walking tests to discriminate fallers from non-fallers [1,10,12,17,22,28]. The SWWT demonstrated good specificity (sensitivity=15%-53%, specificity=70%-97%) for identifying fallers [12,17], and good sensitivity (73%) and specificity (69%) for identifying repeat fallers [17]. However, TUG-manual test (sensitivity=17%, specificity=95%, AUC=0.63) [10], tests from Yang et al. (walking tasks: forward walking at self-selected and maximal speed, backward walking, obstacle course, TUG; secondary tasks: verbal fluency, mental tracking, carrying a glass of water) [28], and working memory tasks during walking [1] were not useful for discriminating fallers from non-fallers.

The criterion-related validity was evaluated by assessing the correlations among DT walking tests. The time taken to complete the TUG test and other walking tests in DT condition yielded moderate-good or good-excellent correlations (Pearson's r=0.66-0.93), while the correlations between the CRR generated from the TUG test and that from other walking tests were weaker (Pearson's r=0.37-0.65) [28]. Convergent validity was assessed by Lee et al. [5], in which a significant difference was found in DT walking performance between the limited community ambulators (i.e., not ambulant outside the home/ambulant as far as the letterbox/ambulant in the immediate environment) and community ambulators (i.e., ambulant in a shopping center and/or places of special interest). Within the group of limited community ambulators, a significant difference in DT walking performance was found between those who had good motor recovery (Brunnstrom stage 5-6) and those who had poor motor recovery (Brunnstrom stage 3-4) [5].

3.5.1.2 Static Sitting

Validity: Convergent validity was assessed in one study [16]. Fair correlation was identified between the variability of body sway during repetitive utterance (Pearson's r=-0.43--0.37) or word generation (r=-0.42) in sitting position and the Barthel ADL index [16].

3.5.2. Parkinson's disease (8 studies)

3.5.2.1 Walking

Reliability: Only two studies have examined the reliability of gait measures in DT condition [2, 13]. These measures (speed, stride length, stride time, swing percentage, cadence, stride width) had good reliability [2], regardless of the type of cognitive tasks used (digit span, auditory Stroop, mobile phone task) (Spearman's rank=0.83-0.95). The reliability was considerably lower for step length asymmetry (Spearman=0.48-0.77), or variability in stride length (Spearman=0.44-0.72) and stride time (Spearman=0.52-0.73) [2]. When DTE values of gait variables were used as outcomes, the reliability was generally lower (ICC=0.19-0.81) [2]. Good internal consistency and moderate-good inter-rater reliability was demonstrated for the item of walking while quoting animal names in the Dynamic Parkinson Gait Scale (DYPAGS) (Cronbach's alpha=0.95; Inter-rater: Kendall's coefficient=0.79, Krippendorff's alpha=0.69, ICC=0.80) [13].

The reliability of cognitive outcomes under DT condition was examined in one study only [2]. Reaction time/response time measures (ICC=0.69-0.82) had better reliability than error rate (Spearman=0.21-0.62). Similarly, poor reliability was reported when the DTE values were used for analysis (reaction time/response time: ICC=0.41, error rate: Spearman=0.09-0.59) [2].

Validity: Other authors explored whether DT walking assessments were useful in identifying fallers among individuals with PD. The time taken to complete the TUG in DT condition (mental tracking: sensitivity=76.5%, specificity=73.7%, AUC=0.82) or (carrying a glass of water: sensitivity=29.6%, specificity=68.4%, AUC=0.78) [26], and the combined motor DTC during the DT walking/box-checking task (sensitivity=71.4%, specificity=77.3%, AUC=0.71) were useful in identifying fallers [27]. In contrast, all outcome variables derived from the DT walking/serial-7-subtractions, DTC of walking speed and DTC of box-checking during the DT walking/box-checking task [27] were not useful in identifying fallers.

The ability of DT assessments to differentiate freezers and non-freezers was also evaluated in one study [26]. The step height during the stepping-in-place task while performing serial-7 subtractions was found to be useful in identifying freezers (sensitivity=70%, specificity=92.9%, AUC=0.86) [26].

The criterion-related validity of DT walking assessments was evaluated in two studies [14,18]. Good correlations were found between the Vitaport Activity Monitor (VAM) and the GAITRite in measuring gait speed, step length and step frequency in DT condition [ICC=0.85-0.98] [18]. The DTC of word generation per second during comfortable-speed walking was also shown to be predictive of impairment and disability level (i.e., convergent validity) [14].

3.5.2.2 Static Standing

Convergent validity of DT standing balance assessment was examined in one study [25]. No significant correlations between DT standing balance ability as indicated by the posturographic center of pressure variables and the clinical balance tests (Berg Balance Scale, Mini-Balance Evaluation System Test) [25].

3.5.3. Mild Cognitive Impairment, Dementia, and Alzheimer's Disease (4 studies)

Reliability: The reliability of the forward walking task [11,19] and 5-meter turn walking test [21] was good, regardless of the difference in cognitive impairment level (ICC=0.81-0.99) [11,19,21]. However, poor reliability was found for the coefficient of variation (CoV) of stride time (ICC=0.34) [11]. The absolute reliability for walking time in a mixed population of MCI and AD (intra-rater, SEM=2.15s, repeatability=5.97s; inter-rater, SEM=0.17s, repeatability=0.47s) was also established by Pettersson et al. [21].

Validity: Convergent validity was assessed in one study [6], in which significant, fair associations between DT forward walking speed and cognitive function (r=0.238-0.395) was found in individuals with MCI [6].

3.5.4. Multiple Sclerosis (3 studies)

Reliability: Good reliability of the DT walking tests was identified, with ICC of gait-related parameters (velocity, cadence, step length, etc.) ranging from 0.80 to 0.97 and SEM%<17% and MDC_{95}%<44% [20].

Validity: A fair correlation was found between the DTC of gait-related parameters and the Physiological Profile Assessment (PPA) fall risk score (r=0.39) [23]. The DTC% of walking time could not significantly predict falls (odds ratio=1.0)

[15].

4. Discussion

Despite different populations and testing protocols used, the overall results showed that the reliability of walking tasks in DT condition was generally good. There was some evidence that the reliability of the variables derived from the secondary cognitive/motor tasks in DT walking assessments and DTE was somewhat lower but this was understudied. The ability of DT walking assessment to distinguish fallers from non-fallers varied, but the sensitivity tended to be lower than the specificity.

4.1.Reliability

The reliability of gait parameters was good (ICC=0.80-0.99) [11-13,18-21], with only a few exceptions. The reliability of the gait parameters was also generally similar regardless of the secondary cognitive/motor tasks used. It is interesting that despite the potential differences in balance and walking performance across the different neurological disorders, the reliability findings were quite consistent. These results were in line with those found in the general elderly population [3], where the gait spatiotemporal parameters in the DT condition also yielded good reliability.

When the DTE values of gait parameters were used as outcomes, the reliability was not as good (ICC=0.19-0.81), although this issue was examined in 3 studies only (two in stroke, one in PD) [2,22,28]. This was also in line with findings in the general elderly population, where poor reliability was typically found for the DTE of gait parameters [3]. In the calculation of DTE, the measurement errors in both the ST and DT conditions were taken into account. Any inconsistencies across trials in each of the task conditions may lead to an inflation of the overall measurement error in DTE, and thus lower reliability values [3]. However, it was reported by Tsang et al. [22] that the reliability of the DTE of the time taken to complete the walking task under DT condition among individuals with chronic stroke was good (Kappa=0.70-0.76) [22]. The relatively high reliability reported in their study may be because the absolute value of DTE was converted into an ordinal scale of 0-2 [22]. This was a very gross measure of DTE only and any difference in DTE across trials may go undetected.

Variability of gait parameters (i.e. coefficients of variation for stride length, step time, etc.) can be useful to identify DT deficits in individuals with neurologic disorders [2,11]. However, only poor-moderate reliability was found for outcomes derived from variability of gait parameters among individuals with PD (Spearman=0.44-0.72) [2]. The reliability was even poorer among individuals with dementia (ICC=0.34) [11]. It was also shown in previous studies on the general older adult population that the reliability was not as satisfactory if the variability in gait parameters was used as an outcome measure [3]. As Hollman et al. [29] proposed, it would require data collected from approximately 220 strides to achieve a magnitude of reliability at 0.9 for an accurate estimation of variability in stride velocity. However, the walkway used in Beauchet et al. was only 10 meters, which was clearly too short a distance for achieving 220 strides [11]. This may account for the inconsistency of the gait variability readings across trials. The number of steps used for analysis in Strouwen et al. [2] was somewhat greater (21 - 125 steps). This may partially explain why higher reliability values were reported in their study, compared with Beauchet et al. [11]. In addition, it cannot be ruled out that the difference in cognitive function may also partly explain the difference in results. The PD patients in Strouwen et al. [2] had no cognitive deficits (MMSE median: 28) whereas those in Beauchet et al. [11] suffered from dementia.

Relatively few studies have assessed the reliability of the outcomes generated from the secondary cognitive and motor tasks under DT condition [2,28]. The principal finding was that the reliability of these parameters was lower than that of the gait parameters. No systematic difference in reliability could be identified across the different cognitive or motor tasks. The findings are in concordance with those in the general older adult population [3]. Many factors such as psychological state, attention, and quality of sleep on the night prior to testing may affect cognitive performance more than gait performance. Cognitive performance may also be prone to a more pronounced learning effect than the walking task [2].

In addition to the relative reliability measures (e.g., ICC values), the absolute reliability (e.g., standard error of the measurement, minimal detectable change) of DT measures was also established in a few studies [2,20,21,28]. These values are essential when interpreting changes in DT performance over time. For example, in the MCI and AD population, a mean change in DT walking time as a group beyond the value of 2.1s would indicate a real change in performance, rather than the errors related to repeated measurements [21].

4.2.Validity Analysis

Several studies assessed convergent validity of the DT walking assessment [5,6,14,16,23,25] and generally, the results were significant. The DT/DTC in walking speed (MS and MCI) [6,23], DTC in word generation during walking (PD) [14], and DT in body sway during static sitting (stroke) [16] had fair correlations with other functional scales [6,14,16,23]. DT walking performance also differed depending on the ambulatory status and motor recovery [5]. To a certain extent, these findings have established the convergent validity of selected DT sitting balance and walking assessments.

Prediction of falls in individuals with neurological conditions is an important topic. Several studies investigated the ability of DT walking to predict falls or distinguish fallers from non-fallers among individuals with stroke, with mixed results [1,10,12,17,22,28]. In particular, the SWWT was used in two studies to predict falls among individuals with stroke, and similar results were identified (i.e. low in sensitivity, high in specificity) [12,17]. It seems that the definition of fallers might have an influence on the value of sensitivity, thus leading to different conclusions. If the SWWT was used to discriminate between recurrent fallers and non-recurrent fallers, rather than the non-fallers and fallers, the sensitivity would be much higher (73% Vs 53%) [17]. The specificity was consistently quite high, however (70%-97% for discriminating between fallers and non-fallers, 69% for distinguishing recurrent fallers from non-recurrent fallers).

The value of DT assessments in distinguishing fallers/non-fallers in PD was investigated in two studies only and the results were mixed [24,27]. Only the combined DTC of the walking/box-checking task could effectively discriminate fallers from non-fallers. Similar to the studies in stroke, the sensitivity (71.4%) reported was lower than specificity (77.3%). The difference in results between sensitivity and specificity values may be related to the fact that the causes of falls are multifactorial. While non-fallers should perform well in DT mobility assessments, fallers may not necessarily have poorer performance in the same DT assessments, as other health or environmental factors may also contribute to their falls. This may partly explain why there was a tendency for the specificity to be higher than the sensitivity.

Research on using DT assessments to predict falls in other neurological conditions is relatively scarce (1 study in MS, no study in MCI or dementia) and therefore inconclusive. Overall, the usefulness of DT balance/walking assessments in predicting falls in individuals with neurological disorders is inconclusive and requires further investigations.

4.3. Limitations of the Studies Reviewed

The studies reviewed had several important limitations. First, of the 23 studies included in this review, very few were rated as good or excellent in terms of methodological quality according to the criteria described in COSMIN (supplementary tables 1-5) [7]. Second, stepping-in-place was used in the testing paradigm in Chomiak et al. [26]. Although both stepping-in-place and walking involve reciprocal, rhythmic movements of the lower extremities, some key spatiotemporal parameters may differ between the two tasks. Indeed, Garcia et al. showed that the single limb support duration of each limb was significantly longer for the stepping-in-place task compared with walking among healthy adults as well as individuals with stroke [30]. Therefore, the findings derived from the stepping-in-place task may not be completely generalizable to the walking task. Third, the reliability of outcomes derived from static/dynamic sitting or standing balance tasks was not examined in any of the selected studies. In addition, the reliability of secondary cognitive/motor task performance was examined in only two studies [2,28], whereas the reliability of the DTE measures was investigated in three studies [2,22,28]. Consequently, solid conclusions on these aspects could not be drawn. Certain cognitive task categories, such as working memory, and discrimination and decision making, are also understudied.

4.4.Limitations of this Systematic Review

There was insufficient data on the validity of DT balance and walking assessments in fall prediction specific to each neurological condition. As a result, meta-analysis could not be performed to estimate the overall odds ratio.

4.5 Recommendations for Future Research

Future good-quality studies should evaluate the reliability of outcomes derived from the static/dynamic sitting and standing balance tasks, secondary cognitive/motor tasks, as well as DTE measures. In general, the psychometric properties of dual-task balance and walking assessments in many neurological disorders are either understudied (e.g., MCI, dementia and MS) or unexplored (e.g., traumatic brain injury). More research on psychometric properties in these neurological populations should warrant further investigations. The multifactorial nature of falls should be considered, and the value of using the DT balance and walking assessments together with other outcomes in predicting falls should be further explored using structural equation modeling.

5. Conclusion

In DT balance and walking assessment for individuals with stroke, PD, MCI, dementia, AD, and MS, parameters derived from the walking tasks can be reliably measured. The psychometric properties of other DT outcomes need to be further investigated.

Acknowledgement

L.Y., M.Y.C.P., and C.Q.H. provided concept/idea/research design. L.Y. provided writing. M.Z.H, L.Y., L.R.L. provided data collection and quality assessment of the study. L.Y. and F.M.H.L. provided data analysis. M.Y.C.P. provided project management. All authors provided consultation (including review of manuscript before submission).

Funding

Mr. Yang was supported by a PhD research studentship provided by the Institute for Disaster Management and Reconstruction, Sichuan University – Hong Kong Polytechnic University (grant number RTHV). Mr. Lam and Ms. Huang were supported by a full-time PhD studentship provided by the Hong Kong Polytechnic University (grant numbers RT5F, RTSA). Mr. Liao and Prof. Pang were supported by the General Research Fund provided by the Research Grants Council (grant number PolyU 5245/11E).

Disclosure of Conflicts of interest

All authors declare no conflict of interest.

Supporting Information

Supplementary Information can be found in the online version of this article:

Supplementary Appendix 1. Example of search strategy.

Supplementary Table 1-5. COSMIN ratings for Box A, B, C, F, H, respectively.

References

[1] Hyndman D, Ashburn A, Yardley L, Stack E. Interference between balance, Gait and cognitive task performance among people with stroke living in the community. Disabil Rehabil 2006; 28: 849-856.

[2] Strouwen C, Molenaar EA, Keus SH, Munks L, Bloem BR, Nieuwboer A. Test-retest reliability of dual-task outcome measures in people with Parkinson disease. Phys Ther 2016; 96: 1276-1286.

[3] Yang L, Liao LR, Lam FM, He CQ, Pang MY. Psychometric properties of dual-task balance assessments for older adults: A systematic review. Maturitas 2015; 80: 359-369.

[4] 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: 289-297.

[5] Lee KB, Kim JH, Lee KS. The relationship between motor recovery and gait velocity during dual tasks in patients with chronic stroke. J Phys Ther Sci 2015; 27: 1173-1176.

[6] Doi T, Shimada H, Makizako H, Tsutsumimoto K, Uemura K, Anan Y, Suzuki T.

Cognitive function and gait speed under normal and dual-task walking among older adults with mild cognitive impairment. BMC Neurology 2014; 14: 67.

[7] 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: 651-657.

[8] Portney LG, Watkins MP. Foundations of Clinical Research. Applications to Practice. 3rd ed. Upper Saddle River, NJ: Pearson Prentice Hall; 2014.

[9] Hosmer DW Jr., Lemeshow S. Applied logistic regression. New York: Wiley; 2000.

[10] Andersson A, Kamwendo K, Seiger A. How to identify potential fallers in a stroke unit: validity indexes of 4 test methods. J Rehabil Med 2006; 38: 186-191.

[11] 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.

[12] Cho KH, Lee HJ, Lee WH. Test-retest reliability of the GAITRite walkway system for the spatio-temporal gait parameters while dual-tasking in post-stroke patients. Disabil Rehabil 2015; 37: 512-516.

[13] Crémers J, Phan Ba R, Delvaux V, Garraux G. Construction and validation of the Dynamic Parkinson Gait Scale (DYPAGS). Parkinsonism Relat Disord 2012; 18: 759-764.

[14] Fuller RL, Van Winkle EP, Anderson KE, Gruber-Baldini AL, Hill T, Zampieri C, Weiner WJ, Shulman LM. Dual task performance in Parkinson's disease: a sensitive predictor of impairment and disability. Parkinsonism Relat Disord 2013; 19: 325-328.

[15] Gunn H, Creanor S, Haas B, Marsden J, Freeman J. Risk factors for falls in multiple sclerosis: an observational study. Mult Scler 2013; 19: 1913-1922.

[16] Harley C, Boyd JE, Cockburn J, Collin C, Haggard P, Wann JP, Wade DT. Disruption of sitting balance after stroke: Influence of spoken output. J Neurol Neurosurg Psychiatry 2006; 77: 674-676.

[17] Hyndman D, Ashburn A. "Stops walking when talking" as a predictor of falls in people with stroke living in the community. J Neurol Neurosurg Psychiatry 2004; 75: 994-997.

[18] Lord S, Rochester L, Baker K, Nieuwboer A. Concurrent validity of accelerometry to measure gait in Parkinsons Disease. Gait Posture 2008; 27: 357-359.

[19] Montero-Odasso M, Casas A, Hansen KT, Bilski P, Gutmanis I, Wells JL, Borrie MJ. Quantitative gait analysis under dual-task in older people with mild cognitive impairment: a reliability study. J Neuroeng Rehabil 2009; 6: 35.

[20] Monticone M, Ambrosini E, Fiorentini R, Rocca B, Liquori V, Pedrocchi A, Ferrante S. Reliability of spatial-temporal gait parameters during dual-task interference in people with multiple sclerosis. A cross-sectional study. Gait Posture 2014; 40: 715-718.

[21] Pettersson AF, Olsson E, Wahlund L-O. Effect of divided attention on gait in subjects with and without cognitive impairment. J Geriatr Psychiat Neurol 2007; 20: 58-62.

[22] Tsang CS, Liao LR, Chung RC, Pang MY. Psychometric properties of the Mini-Balance Evaluation Systems Test (Mini-BESTest) in community-dwelling individuals with chronic stroke. Phys Ther 2013; 93: 1102-115.

[23] Wajda DA, Motl RW, Sosnoff JJ. Dual task cost of walking is related to fall risk in persons with multiple sclerosis. J Neurol Sci 2013; 335: 160-163.

[24] Vance RC, Healy DG, Galvin R, French HP. Dual tasking with the timed "up & go" test improves detection of risk of falls in people with Parkinson disease. Phys Ther 2015; 95: 95-102.

[25] Barbosa AF, Souza Cde O, Chen J, Francato DV, Caromano FA, Chien HF, Barbosa ER, Greve JM, Voss MC. The competition with a concurrent cognitive task affects posturographic measures in patients with Parkinson disease. Arq Neuropsiquiatr 2015; 73:906-912.

[26] Chomiak T, Pereira FV, Meyer N, de Bruin N, Derwent L, Luan K, et al. A new quantitative method for evaluating freezing of gait and dual-attention task deficits in Parkinson's disease. J Neural Transm (Vienna) 2015; 122: 1523-1531.

[27] Heinzel S, Maechtel M, Hasmann SE, Hobert MA, Heger T, Berg D, Maetzler W. Motor dual-tasking deficits predict falls in Parkinson's disease: A prospective study. Parkinsonism Relat Disord 2016. DOI: 10.1016/j.parkreldis.2016.03.007.

[28] Yang L, He C, Pang MY. Reliability and validity of dual-task mobility assessments in people with chronic stroke. PLoS ONE 2016; 11: e0147833.

[29] 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: 23-28.

[30] Garcia RK, Nelson AJ, Ling W, van Olden C. Comparing stepping-in-place and gait ability in adults with and without hemiplegia. Arch Phys Med Rehabil 2001; 82: 36-42.

FIGURE LEGENDS

Figure 1. Flow diagram.

The diagram illustrates the flow of information through the different phases of the systematic review. Twenty-three articles were included in this review.