Measures of balance and falls risk prediction in people with Parkinson's disease: a systematic review of psychometric properties

Winser, Stanley J.; Kannan, Priya; Bello, Umar Muhhamad; Whitney, Susan L.

Published in:
Clinical Rehabilitation

DOI:
10.1177/0269215519877498

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Original Article

CLINICAL REHABILITATION

Clinical Rehabilitation
2019, Vol. 33(12) 1949–1962

DOI: 10.1177/0269215519877498
journals.sagepub.com/home/cre
SAGE

Measures of balance and falls risk prediction in people with Parkinson's disease: a systematic review of psychometric properties

Stanley J Winser[1], Priya Kannan[1], Umar Muhhamad Bello[1] and Susan L Whitney[2]

Abstract

Objective: To investigate the psychometric properties of measures of balance and falls risk prediction in people with Parkinson's disease (PD).

Data sources: PubMed, Embase, CINAHL, Ovid Medline, Scopus, and Web of Science were searched from inception to August 2019.

Review method: Studies testing psychometric properties of measures of balance and falls risk prediction in PD were included. The four-point COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) assessed quality.

Results: Eighty studies testing 68 outcome measures were reviewed; 43 measures assessed balance, 9 assessed falls risk prediction, and 16 assessed both. The measures with robust psychometric estimation with acceptable properties were the (1) Mini-Balance Evaluation Systems Test (Mini-BEST), (2) Berg Balance Scale, (3) Timed Up and Go test, (4) Falls Efficacy Scale International, and (5) Activities-Specific Balance Confidence scale. These measures assess balance and falls risk prediction at the body, structure and function level, falls risk and balance, and falls risk at the activity level. The motor examination of the Unified Parkinson's Disease Rating Scale (UPDRS-ME) with robust psychometric analysis is a condition-specific measure with acceptable properties. Except the UPDRS-ME and Mini-BESTest, the responsiveness of the other four measures has yet to be established.

Conclusion: Six of the 68 outcome measures have strong psychometric properties for the assessment of balance and falls risk prediction in PD. Measures assessing balance and falls risk prediction at the participatory level are limited in number with a lack of psychometric validation.

Keywords

Parkinson's disease, balance and falls, reliability, validity, falls risk

Received: 19 May 2019; accepted: 29 August 2019

[1]Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Kowloon, Hong Kong
[2]Department of Physical Therapy, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, PA, USA

Corresponding author:
Stanley J Winser, Suite ST514, Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Kowloon, 11 Yuk Chio Rd, Hung Hom, Hong Kong.
Email: Stanley.j.winser@polyu.edu.hk

Introduction

People with Parkinson's disease (PD) are at an increased risk of falling, and measures of balance and measures of falls risk prediction are required. There are many, but there is no review of the evidence as to which are better or worse. It is of paramount importance to adopt assessment tools with sound psychometric properties to ensure accuracy and reproducibility in assessing balance and predicting falls in persons with PD.

In PD, a recent critical review by the Movement Disorders Society Task Force reported a set of "recommended," "suggested," and "listed" measures of balance, gait, and falls.[1] The authors of that review selected common measures of balance, gait, and falls risk prediction and recommended measures based on the findings of psychometric properties, validation research performed in samples of individuals with PD and if data were available for the outcome measures' use in clinical studies beyond the outcome measures developer's group. However, the recommendations are restricted to those outcome measures that do not need extra tools for administration, that is, those that could be administered at the bedside. In addition, their included studies were not systematically pooled from specific databases, which might have limited the robust inclusion of published studies in the research area. Furthermore, their recommendations were not based on the International Classification of Functioning, Disability and Health model, that is, they did not take into account (1) body, structure, and function; (2) activity; or (3) participation levels of assessment for estimating balance and falls risk prediction.[2]

Given these considerations, the objectives of this review are to perform a systematic review of the psychometric properties of measures of falls and falls risk in individuals with PD in order to (1) identify those measures with the strongest psychometric properties; (2) classify the available outcome measures into that that assess balance and falls risk at the (a) body, structure, and function, (b) activity, or (c) participatory levels; and (3) discuss the implications of the findings for clinical practice and future research that could provide additional testing of the psychometric properties of the existing measures for assessing balance and falls risk in individuals with PD.

Methods

Search strategy

The following electronic databases were searched from inception to August 2019: PubMed, Embase, CINAHL, Ovid Medline, Scopus, and Web of Science. The search terms were constructed using the following four themes: PD, psychometric properties, balance and falls, and outcome measures. Related terms were combined using the Boolean "OR"; all the themes were then combined using the Boolean "AND" (Supplemental Appendix 1 reports the search terms used for the database EMBASE). To ensure a thorough search, 18 common measures of balance and falls risk prediction utilized for persons with PD were included in the search theme "outcome measure" and were combined using the Boolean "OR."

Studies that fulfilled the following criteria were included: (1) assessed one or more of the following psychometric properties: internal consistency, reliability, measurement error, content validity, face validity, structural validity, hypothesis testing, cross-cultural validity, criterion validity, or responsiveness; (2) psychometric analysis was done among people with PD; (3) outcome measures including clinical and laboratory-based assessment of balance or falls risk prediction or both; and (4) studies published in English language. Studies were excluded if they were (1) conference abstracts, (2) psychometric property testing protocols, or (3) studies testing the psychometric property of gait analysis, freezing of gait, or other non-motor symptoms associated with PD. In this review, we define reliability as the extent to which the scores of the outcome measure are reproducible when the assessment is repeated by the same or different examiner,[3] validity as the extent to which the instrument measures what it is intended to measure[3] and responsiveness as the ability of the outcome measure to detect changes over time.[3]

Screening, data extraction, and categorization

All the retrieved studies were subject to a four-level screening process that included duplicate removal, title, abstract, and full-text screening. Two authors (U.M.B. and S.J.W.) were involved in the screening process. The following data were extracted from the included studies: title, objectives, outcome measure studied, psychometric properties tested, and the reported findings. Retrieved outcome measures were grouped into either measures of balance or measures of falls risk prediction. We used the International Classification of Functioning, Disability and Health model to further categorize the measures of balance and falls risk prediction. Three reviewers (S.J.W., P.K., and S.L.W.) classified the measures according to the level of assessment using the International Classification of Functioning, Disability and Health model into one of the following levels: (1) body structure and function, (2) activity, or (3) participation.[2] We used the recommendations by the Parkinson Edge Outcome Measures Taskforce (http://www.neuropt.org/docs/default-source/parkinson-edge/single-measure-detailed-ratings820e33a5390366a68a96ff00001fc240.pdf?sfvrsn=ba0d5543_0) for this categorization. For the outcome measures that were not listed by the Taskforce, the three reviewers independently assessed the outcome measure, and any discrepancies between reviewers were discussed until consensus was reached on the category of the outcome measure.

Quality appraisal

The methodological quality of the psychometric properties and the level of evidence of the measures of balance and falls risk prediction were evaluated using the four-point COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN).[3] The COSMIN is a reliable and valid quality appraisal tool for systematic reviews of psychometric properties of outcome measures.[3,4] Based on the scores obtained, the psychometric property was rated as "excellent," "good," "fair," or "poor." Studies were not excluded based on quality. The methodological quality of the psychometric properties of the identified outcome measures was completed independently by two reviewers (S.J.W. and U.M.B). Discrepancies were resolved by discussion between the reviewers. A third reviewer (S.L.W) was consulted for any unresolved discrepancies.

We used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to report the findings of this systematic review.

Results

The search identified 1625 studies, of which 80 studies were eligible for inclusion in this systematic review. The included studies yielded 68 outcome measures assessing balance or falls risk prediction or both. Figure 1 illustrates the flow of data search and screening. Supplemental Appendix 2 presents a summary of the included studies with references. Table 1 lists the identified measures assessing balance, falls risk prediction, and both balance and falls risk prediction corresponding to the level of assessment according to the International Classification of Functioning, Disability and Health model. This review identified 43 measures assessing balance, 9 assessing falls risk prediction and 16 assessing both in individuals with PD. Fourteen measures assessed balance and/or falls risk at the body, structure, and function level; 50 at the activity level; and 4 at the participatory level. Among the identified measures, 14 were condition-specific, and the remaining 54 were generic measures of balance or falls risk. Supplemental Appendices 3–5 present lists of the measures assessing balance and falls risk at the (1) body, structure, and function level; (2) activity level; and (3) participatory levels, respectively, as well as their COSMIN quality scores.

Measures of balance for PD

The psychometric properties of the following measures have been evaluated extensively in samples of individuals with PD: The Balance Evaluation Systems Test (BESTest),[5–7] Mini-BESTest,[6–18] Sensory Organization Test,[19–21] Berg Balance Scale (BBS),[14,17,18,22–30] Forward Functional Reach,[24,29,31–34] Timed

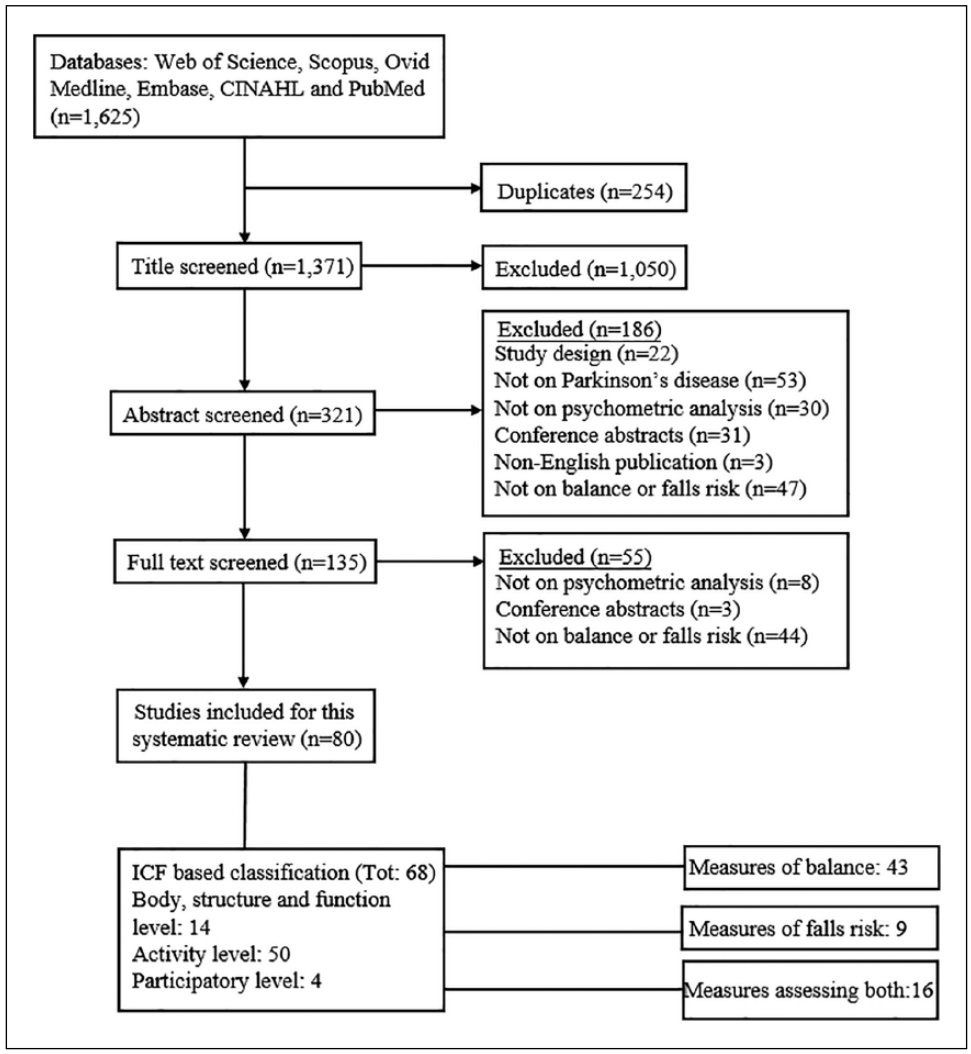

Figure 1. Screening of studies for inclusion.

Up and Go (TUG) test,[16,17,23–25,29–31,35–41] Motor examination of the Unified Parkinson's Disease Rating Scale (UPDRS-ME),[12,13,17,23,27,29,30,31,36,42,43] and the Activities-specific Balance Confidence[10,29,30,35,44–46] scale. Among these measures, the Mini-BESTest assessing balance at the body structure and function level has good to excellent inter-rater reliability (intraclass correlation coefficient (ICC) $>$ 0.95),[17] test re-test reliability (ICC $>$ 0.95),[17] and internal consistency (Cronbach's alpha $=$ 0.87).[15] The COSMIN quality of these estimates was poor; however, in terms of validity, Rasch analysis reporting adequate structural validity,[8] adequate predictive validity,[12] discriminant validity,[16] concurrent validity,[17] and convergent validity[11] have been reported. The COSMIN quality of validity estimates for the Mini-BESTest is good. Good COSMIN quality psychometric estimation found the Mini-BESTest responsive to balance related changes among people with PD at 6 and 12 months.[12]

Among the generic measures assessing balance at the activity level, the Activities-specific Balance

Table 1. List of the identified measures of balance and falls risk prediction for people with Parkinson's disease reported alphabetically.

ICF classification	Identified outcome measures		
Body structure and function level Total outcome measures = 14	**Measures of balance (*n* = 10)**		
	Balance master	Limits of Stability Test (LOS)	Quantitative Posturography (QP)
	Center of Mass (COM) using mobile device.	Maximal Balance Range (MBR)	Sway on floor and foam
	GAITRite	Motor Control Test (MCT)	
	Instrumented Timed Up and Go (i-TUG)	Nintendo Wii©	
	Measures of balance and falls risk prediction (*n* = 4)		
	Balance Evaluation System Test (BESTest)	Mini-Balance Evaluation Systems Test (Mini-BESTest)	Sensory Organization Test (SOT)
	Modified Clinical test for the Sensory Integration of Balance (mCTSIB)		
Activity level Total outcome measures = 50	**Measures of balance (*n* = 33)**		
	Ambulatory capacity measure (ACM)	Movement Disorder Society–Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS)	Six-Spot Step Test
	Backward Functional Reach Test (BFR)	Nutt et al. protocol	Short Parkinson Evaluation Scale (SPES)
	Bäckstrand Dahlberg Liljenäs Balance Scale (BDL)	Parkinson's Disease Composite Scale (PDCS)	30 seconds chair stand
	Bed mobility test (BMT)	Pastor et al. protocol	Timed Up and Go Assessment of Biomedical Strategies (TUG-ABS)
	Bloem et al. protocol	Postural Change Scale (PCS)	Tinetti Mobility Test (TMT)
	Comfortable 10-Meters walk Test (CMT)	Postural Stability Score (PSS)	Unified Parkinson's Disease Rating Scale–Motor Examination (UPDRS-ME)
	Coordinated Stability Task (CST)	Pull Test	Unified Parkinson's Disease Rating Scale (UPDRS)

(Continued)

Table 1. (Continued)

ICF classification	Identified outcome measures		
	Choice Stepping Reaction Time (CSRT)	Push and Release Test (P&RT)	Unified Parkinson's Disease Rating Scale (UPDRS) items 27–29
	Fast 10 Meters Walk Test (FWT)	Reach test	Visual Analog Scale (VAS)
	Fullerton Advanced Balance scale (FAB)	Single-Leg-Stance-Test (SLST)	30 seconds sit-to-stand test (30sSTS)
	Gait speed	Single-Leg-Stance (SLS) ± acutely induced dizziness	360° turn test
	Measures of falls risk prediction (*n* = 5)		
	Falls Efficacy Scale (FES)	Fall-related Activity Classification (FRAC)	Three-steps falls prediction model
	Falls Efficacy Scale–International (FES-I)	Functional Gait Assessment (FGA)	
	Measures of balance and falls risk prediction (*n* = 12)		
	Activities-specific Balance Confidence (ABC)	Forward Functional Reach (FFR)	Optoelectronic system Timed Up and Go (Opto-TUG)
	Berg Balance Scale (BBS)	Four Square Step Test (FSST)	Timed Up and Go (TUG)
	Dynamic Gait Index (DGI)	Freezing of Gait Questionnaire (FOGQ)	5-item Activities-specific Balance Confidence (ABC-5I)
	Five Times Sit-to-Stand Test (FTSTS)	Gait and Balance Scale (GABS)	6-item Activities-specific Balance Confidence (ABC-6p)
Participatory level Total outcome measures = 4	**Measures of falls risk prediction (*n* = 4)**		
	Modified Short version of Survey of Activities and Fear of Falling in the Elderly (mSAFFE)	Survey of Activities and Fear of Falling in the Elderly (SAFFE)	
	Rapid assessment of postural instability questionnaire (RAPID	Timed Up and Go Cognition (CTUG)	

ICF: International Classification of Functioning, Disability and Health.

Confidence, BBS, and the TUG test were found to be reliable[17,22,28,29,31,35,39,40,41,44] and valid.[11,14,16,17,22–28,30,31,35–38,44,45] The COSMIN quality of both of these estimates ranged from poor and good. The responsiveness supporting its use has been reported for the TUG test[47] but not for the BBS and the Activites-specific Balance Confidence. Among the condition-specific measures assessing balance at the activity level, the UPDRS-ME was found to be reliable,[29,42] valid,[12,13,17,27,30,31,36,43] and responsive to change.[12] The COSMIN quality scores of these estimates ranged between poor[23,27,42] and excellent.[43] Most of the studies testing the UPDRS-ME used the measure as a comparator to establish the psychometric properties of other generic measures of balance. Items 26 through to 31 (6 items) and item 13 on falling of the UPDRS-ME are relevant to the assessment of balance and falls risk. However, none of the identified studies tested the psychometric properties of these selected items. The Pull test and the Push and Release Test assessing balance at the body, structure, and function level was found to have acceptable reliability[45,48] and validity.[32] These estimates arrive from three studies of either poor[45,48] or good[32] COSMIN quality. One study of fair[37] COSMIN quality reported the responsiveness of the Push and Release test as the difference between the test performance between ON and OFF phase following medication. The Push and Release test was found to have a significant difference in scores between ON and OFF phase following medication.

Measures of falls risk prediction for PD

The BESTest,[6,49,50] Sensory Organization Test,[51] and the Mini-BESTest[8,9,49,51] assessing the falls risk prediction at the body structure and function level were commonly subject to psychometric analysis. One low COSMIN quality study supporting adequate reliability[6] and two good COSMIN quality studies supporting adequate predictive validity[49] and discriminant validity[6] were found for the BESTest. Scores less than 69% on the BESTest were found to be 84% sensitive and 76% specific in discriminating between fallers and non-fallers,[6] while scores less than 21% for the Mini-BESTest were found to be 63% sensitive and 100% specific.[51]

The Activities-specific Balance Confidence, BBS, Functional Gait Assessment, and the TUG test have been tested extensively for falls risk prediction at the activity level. The Activities-specific Balance Confidence was found to be reliable[52] and valid[6,52,53] in assessing falls risk. One excellent COSMIN quality study reported Activities-specific Balance Confidence score of $\leqslant$55% as 71% sensitive and 62% specific to discriminate between non-recurrent and recurrent fallers.[53] One low COSMIN quality study supporting reliability (inter-rater reliability ICC = 0.95 and test re-test reliability, ICC = 0.79)[6] of the BBS and four low to good COSMIN quality studies supporting validity (construct,[54] discriminant,[53] predictive,[49] and concurrent)[55] report the BBS as an efficient falls risk-assessing tool in PD. Scores $\leqslant$47 on the BBS had 72% sensitivity and 75% specificity in discriminating fallers and non-fallers.[6] One good COSMIN quality study reported the Functional Gait Assessment inferior to the BBS, BESTest, and Mini-BESTest in predicting falls at 12 months.[49] No reports were found testing the reliability of falls risk assessment using the TUG test; however, the construct[54,56] and discriminant validity[53] was supported by three good COSMIN quality studies and one poor quality study.[55]

The modified TUG test called the dual-task or the cognitive TUG test assessing falls risk at a participatory level is found to have moderate test–retest reliability (ICC = 0.55)[40] and acceptable concurrent validity to assess the cognitive–motor interaction while walking.[25] The Falls Efficacy Scale[56] and Falls Efficacy Scale–International[52,53,57] testing falls risk at the activity level and the Survey of Activities and fear of Falling in the Elderly[56] assessing falls risk at the participatory level have been commonly tested for psychometric properties. The Falls Efficacy Scale–International was found to be reliable,[52,57] valid,[57] and able to discriminate between people who were afraid of falls, avoided activities, and experienced falls.[57] One good COSMIN quality study reported excellent test–retest reliability (ICC = 0.92), internal consistency (α = 0.95), adequate construct validity,

and insignificant floor and ceiling effect[56] for the Survey of Activities and fear of Falling in the Elderly scale. Two good and one excellent COSMIN quality studies reported good test–retest reliability (ICC $>$ 0.80),[52] internal consistency ($\alpha = 0.96$),[57] and adequate convergent[57] and discriminant validity (non-recurrent fallers versus recurrent fallers).[53]

The Freezing of Gait Questionnaire assessing falls risk at the activity level and the rapid assessment of postural instability questionnaire assessing falls risk at the participatory level are condition-specific measures assessing of falls risk. The Freezing of Gait Questionnaire assesses the severity of freezing of gait unrelated to falls in people with PD. Based on the available literature, both the Freezing of Gait Questionnaire and Rapid assessment of postural instability questionnaire do not have sufficient psychometric property evaluation to recommend their use in assessing falls risk in this population.

Discussion

This systematic review identified the following measures of balance and falls risk as psychometrically sound: (1) the Mini-BESTest assessing balance and falls risk prediction at the body, structure, and function level and (2) the Falls Efficacy Scale–International assessing falls risk and the Activites-specific Balance confidence, BBS, and the TUG Test assessing balance and falls risk at the activity level. However, despite these positive findings, a strong recommendation on the use of the Activites-specific Balance confidence, BBS, and Falls Efficacy Scale–International cannot yet be made, as the responsiveness of these measures has yet to be established in people with PD. We identified the UPDRS-ME as the only condition-specific tool that has been tested and found to have strong psychometric properties. Current evidence on two other condition-specific measures assessing balance at the body, structure, and function level; the Pull test and Push and Release suggest adequate reliability and validity; however, future research is needed to estimate the responsiveness in order to make firm recommendations on their use.

Most of the measures assessing balance and falls risk prediction at the body, structure, and function level used a laboratory-based or sophisticated instrument. These instruments assessed the ability to shift the center of gravity,[40,58] center of mass,[21] spatiotemporal parameters while walking,[58] and sensory integration[19,20,51] to quantify balance and/or falls risk. We were not able to draw conclusions about the psychometric qualities of or make recommendations regarding the use of these instrumented assessment procedures because they lack evaluation of their psychometric properties. In addition, the use of sophisticated instruments for assessing balance and falls risk has limited clinical utility because they are expensive, such instrumentation is not commonly available in most clinics. Thus, clinic-based or bed-side assessments using these equipment on a routine basis is often not possible. However, for research, it is acknowledged that the use of sophisticated instruments can provide useful information on subtle changes of balance that could not be identified by bed-side clinical tools.

The Mini-BESTest, BBS, TUG Test, Falls Efficacy Scale–International, Activites-specific Balance Confidence, and the UPDRS-ME have the strongest psychometric properties. Moreover, they are brief, are easy to administer, have no cost, and do not require specialized training for the assessor. This review found two condition-specific measures assessing balance and falls risk at the body, structure, and function level (The Pull test and Push and Release test). The available evidence supports adequate reliability and validity; however, there is a lack of estimation of responsiveness. The responsiveness of the Pull test to differentiate fallers from non-fallers has been reported between ON and OFF medication on the same day.[37] However, the disease being progressive, a prospective assessment after a period of time is required to understand the tools ability to pick changes over time.

Among the measures that assessed balance and falls risk at the activity level, the BBS, TUG Test, Falls Efficacy Scale–International, and the Activities-specific Balance Confidence appeared to have the strongest psychometric properties, with high reliability. However, the quality of most of the

reliability estimates for the Activities-specific Balance Confidence,[35,46,52] the BBS,[6,17,22,28] and the TUG test[31,35,39,41] was rated poor according to the COSMIN. The small sample size was one of the common reasons for rating the findings as having low quality. In this review, we used the "worst score counts" algorithm for rating the overall quality of the psychometric properties, as recommended by the COSMIN group.[59] An alternative method of deriving the overall scores from a COSMIN assessment was to calculate the mean score of each section.[59] However, this method was not considered for our review as there was a possibility that major methodological flaws could be compensated by high scores on other aspects of the study design. We strongly recommend that investigators evaluating the psychometric properties of these measures use a sample size of at least 50 or ideally more to allow for high-quality estimates, based on COSMIN standards.[59]

This systematic review did not find estimates of the responsiveness for the BBS, Falls Efficacy Scale–International, and the Activities-specific Balance Confidence scales among people with PD. PD is described as a chronic and progressive disorder; therefore, a measure that is responsive to change over time is needed for research in this area. We are therefore unable to make a firm recommendation on the use of these three measures due to the lack of responsiveness estimates. We recommend future studies to examine the responsiveness of the BBS, Falls Efficacy Scale–International and the Activities-specific Balance Confidence at 6 months or 12 months to allow for strong recommendations. The available evidence for the responsiveness of the TUG test is based on a preassessment and postassessment following eight weeks of group physiotherapy intervention.[47] A longer follow-up assessment to capture the natural progress of the disease is required to make a firm conclusion on the responsiveness of the TUG test.

Fitzpatrick et al.[60] recommend an appropriate set of outcome measures should have one condition-specific measure and a generic measure for assessing a given domain. A condition-specific measure identifies changes that are in close relation or "proximal" to the disease; such postural abnormalities (item 28 of the UPDRS-ME) in PD and the generic measure pick changes that are slightly less proximal or "distal" to the health condition,[61] such as ability to stand unsupported (item 1 of the BBS).

Based on the findings from this review, we recommend six measures, of which only one is condition-specific (UPDRS-ME) and five are generic measures of balance and falls risk for use in assessing these domains in individuals with PD. In light with Fitzpatrick et al.'s[60] recommendation, we propose the use of a combination of one generic and one condition-specific assessment for balance and falls risk prediction. For assessing balance and falls risk prediction at the body structure and function level, a combination of Mini-BESTest and Push and Release test of the UPDRS-ME (item 30) might be considered while a combination of Activities-specific Balance Confidence and/or BBS and/or TUG Test and UPDRS-ME could be adopted for assessing balance at the activity level. A combination of Falls Efficacy Scale–International and UPDRS-ME could be adopted for assessing falls risk prediction at the activity level. The UPDRS-ME has 14 items, with higher scores indicating more motor impairment.

Among the 14 items of the UPDRS-ME, arising from chair (item 27), posture (item 28), gait (item 29), postural stability (item 30), and body bradykinesia and hypokinesia (item 31) plus item 12 under "activities of daily living" are closely related to the domain of balance and falls risk. The utility of using the remaining nine items of the UPDRS-ME to quantify balance and falls risk might be questioned. We recommend that future studies estimate the psychometric properties of the listed items of the UPDRS-ME and determine the validity of using the items specific to balance and falls risk to compute a score, rather than use the total motor examination score.

The Movement Disorders task force recommends the use of the Postural Instability and Gait Difficulty (PIGD), a subscale of the UPDRS, as a measure of balance and gait stability.[1] The PIGD comprises five items from the UPDRS (items 13–15, 29, and 30) assessing falls, freezing, walking ability, and postural stability. Higher scores indicate

greater balance and gait impairment severity. None of the included studies of the current systematic review reported the psychometric properties of the Postural Instability and Gait Difficulty scale. The lack of such studies may be due to our inclusion criteria, as we restricted our search to measures of balance and falls risk related to walking and freezing only. However, the Postural Instability and Gait Difficulty is likely to assess the risk of falls. Therefore, future studies are recommended to assess whether the Postural Instability and Gait Difficulty can discriminate between individuals with a history of falling frequently and occasionally and those who do not have a history falling. Our systematic review identified one study that evaluated the validity of the subscores of items 27–29 of the UPDRS.[32] However, their study recommended the use of a combination of one-leg stance test, pull test, and functional reach test along with items of 27–29 of the UPDRS for optimal assessment of postural stability. Therefore, using items 27–29 of the UPDRS alone is not recommended as a measure of balance.

In summary, we have the following recommendations for future research: (1) There is a need to establish the responsiveness of the Activites-specific Balance Confidence, BBS, Falls Efficacy Scale–International, TUG, Pull test and Push and Release Test in people with PD. (2) Future research on psychometric analysis is recommended to use a sample size of at least 50 or ideally more to allow for high-quality estimates, based on COSMIN standards. (3) There is a need to conduct psychometric analysis of selected items relevant to balance and falls risk prediction among the UPDRS-ME scale to reduce the time spent on assessing balance and falls risk prediction in people with PD. (4) Future studies are recommended to develop or psychometrically validate the available participatory-level outcome measures for people with PD.

Study strengths and limitations

To our knowledge, this is the first systematic review evaluating the psychometric properties of measures of balance and falls risk in people with PD. It has a number of strengths. First, we adopted a systematic search to explore all measures subject to psychometric property testing in people with PD. Second, we used the COSMIN, a valid tool for rating the methodological quality of the included studies. Finally, we included all measures (including both clinic-based and laboratory-based) that assessed balance and falls risk prediction.

However, the study also has a number of important limitations that should be considered when interpreting the results. First, we did not include conference abstracts and non-English studies in the review. Second, all of the included studies recruited participants with mild or moderate PD severity. We are therefore unable to determine the extent to which the findings generalize to samples of patients with severe PD since the severity of PD, including balance levels and falls risk can vary considerably as the disease progresses.[62] In addition, functional losses in gait and balance appear to occur differentially across the different stages of PD progression.[62] Thus, although it is possible that a particular measure may be more or less reliable and valid in individuals at different stages of PD, we were unable to determine the effects of progression stage on the psychometric variables studied. Finally, we did not include any randomized controlled trials in this review. Such studies can provide information on the minimal clinically important differences or minimal detectable changes in measures. However, randomized controlled trials are not the only source for estimating minimal clinically important difference, and we were able to gather and report information for these values in the current article using other statistical methods.

Clinical messages

- When assessing balance in people with PD, Mini-BESTest, and Push and Release test are best at the body level.
- Activities-specific Balance Confidence, BBS, TUG test and the UPDRS-ME are best at the activity level.
- Falls Efficacy Scale-International and UPDRS-ME are best to predict falls risk.

Authors' note

This study was performed at the Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong.

Acknowledgements

The team of authors would like to acknowledge our Research Assistant Mr Kwan Wills for his assistance with data entry and proof reading.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: The open access for this systematic review is supported by the 'Start-up fund' for early careers, The Hong Kong Polytechnic University, Hong Kong.

ORCID iDs

Stanley J Winser https://orcid.org/0000-0002-8766-3688
Priya Kannan https://orcid.org/0000-0003-2583-9614
Umar Muhhamad Bello https://orcid.org/0000-0001-8535-4163

Supplemental material

Supplemental material for this article is available online.

References

1. Bloem BR, Marinus J, Almeida Q, et al. Measurement instruments to assess posture, gait, and balance in Parkinson's disease: critique and recommendations. *Mov Disord* 2016; 31(9): 1342–1355.
2. WHO. International Classification of Functioning, Disability and Health, 2001, https://www.who.int/classifications/icf/en/
3. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; 63(7): 737–745.
4. Winser SJ, Smith CM, Hale LA, et al. COSMIN for quality rating systematic reviews on psychometric properties. *Phys Ther Rev* 2015; 20(2): 132–134.
5. Duncan RP, Cavanaugh JT, Earhart GM, et al. External validation of a simple clinical tool used to predict falls in people with Parkinson disease. *Parkinsonism Relat Disord* 2015; 21(8): 960–963.
6. Leddy AL, Crowner BE and Earhart GM. Utility of the Mini-BESTest, BESTest, and BESTest sections for balance assessments in individuals with Parkinson disease. *J Neurol Phys Ther* 2011; 35(2): 90–97.
7. Maia AlC, Rodrigues-de-Paula Ft, Magalhães LC, et al. Cross-cultural adaptation and analysis of the psychometric properties of the Balance Evaluation Systems Test and MiniBESTest in the elderly and individuals with Parkinson's disease: application of the Rasch model. *Braz J Phys Ther* 2013; 17(3): 195–217.
8. Benka Wallén M, Sorjonen K, Löfgren N, et al. Structural validity of the Mini-Balance Evaluation Systems Test (Mini-BESTest) in people with mild to moderate Parkinson disease. *Phys Ther* 2016; 96(11): 1799–1806.
9. Bergstrom M, Lenholm E and Franzen E. Translation and validation of the Swedish version of the mini-BESTest in subjects with Parkinson's disease or stroke: a pilot study. *Physiother Theory Pract* 2012; 28(7): 509–514.
10. Combs SA, Diehl MD, Filip J, et al. Short-distance walking speed tests in people with Parkinson disease: reliability, responsiveness, and validity. *Gait Posture* 2014; 39(2): 784–788.
11. Duncan RP, Leddy AL and Earhart GM. Five times sit to stand test performance in Parkinson's disease. *Arch Phys Med Rehabil* 2011; 92(9): 1431–1436.
12. Duncan RP, Leddy AL, Cavanaugh JT, et al. Detecting and predicting balance decline in Parkinson disease: a prospective cohort study. *J Parkinsons Dis* 2015; 5(1): 131–139.
13. Duncan RP and Earhart GM. Four Square Step Test performance in people with Parkinson disease. *J Neurol Phys Ther* 2013; 37(1): 2–8.
14. King LA, Priest KC, Salarian A, et al. Comparing the Mini-BESTest with the Berg Balance Scale to evaluate balance disorders in Parkinson's disease. *Parkinsons Dis* 2012; 2012: 375419.
15. Lofgren N, Lenholm E, Conradsson D, et al. The Mini-BESTest: a clinically reproducible tool for balance evaluations in mild to moderate Parkinson's disease? *BMC Neurol* 2014; 14: 235.
16. Löfgren N, Benka Wallén M, Sorjonen K, et al. Investigating the Mini-BESTest's construct validity in elderly with Parkinson's disease. *Acta Neurol Scand* 2017; 135(6): 614–621.
17. Schlenstedt C, Brombacher S, Hartwigsen G, et al. Comparing the Fullerton Advanced Balance scale with the Mini-BESTest and Berg Balance Scale to assess postural control in patients with Parkinson disease. *Arch Phys Med Rehabil* 2015; 96(2): 218–225.
18. Schlenstedt C, Brombacher S, Hartwigsen G, et al. Comparison of the Fullerton Advanced Balance Scale, Mini-BESTest, and Berg Balance Scale to predict falls in Parkinson disease. *Phys Ther* 2016; 96(4): 494–501.

19. Freeman L, Gera G, Horak FB, et al. Instrumented test of sensory integration for balance: a validation study. *J Geriatr Phys Ther* 2018; 41(2): 77–84.
20. Harro CC, Marquis A, Piper N, et al. Reliability and validity of force platform measures of balance impairment in individuals with Parkinson disease. *Phys Ther* 2016; 96(12): 1955–1964.
21. Ozinga SJ, Linder SM and Alberts JL. Use of mobile device accelerometry to enhance evaluation of postural instability in Parkinson disease. *Arch Phys Med Rehabil* 2017; 98(4): 649–658.
22. Babaei-Ghazani A, Mohammadi H, Shahidi G, et al. Reliability and validity of the Persian translation of Berg Balance Scale in Parkinson disease. *Aging Clin Exp Res* 2017; 29(5): 857–862.
23. Claesson IM, Grooten WJ, Lökk J, et al. Assessing postural balance in early Parkinson's Disease: validity of the BDL balance scale. *Physiother Theory Pract* 2017; 33(6): 490–496.
24. Franchignoni F, Martignoni E, Ferriero G, et al. Balance and fear of falling in Parkinson's disease. *Parkinsonism Relat Disord* 2005; 11(7): 427–433.
25. McKee KE and Hackney ME. The Four Square Step Test in individuals with Parkinson's disease: association with executive function and comparison with older adults. *NeuroRehabilitation* 2014; 35(2): 279–289.
26. La Porta F, Giordano A, Caselli S, et al. Is the Berg Balance Scale an effective tool for the measurement of early postural control impairments in patients with Parkinson's disease? evidence from Rasch analysis. *Eur J Phys Rehabil Med* 2015; 51(6): 705–716.
27. Qutubuddin AA, Pegg PO, Cifu DX, et al. Validating the Berg Balance Scale for patients with Parkinson's disease: a key to rehabilitation evaluation. *Arch Phys Med Rehabil* 2005; 86(4): 789–792.
28. Scalzo PL, Nova IC, Perracini MR, et al. Validation of the Brazilian version of the Berg Balance Scale for patients with Parkinson's disease. *Arq Neuropsiquiatr* 2009; 67(3B): 831–835.
29. Steffen T and Seney M. Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-item short-form health survey, and the unified Parkinson disease rating scale in people with Parkinsonism (*Phys Ther* 2008; 88: 733–746). *Phys Ther* 2008; 90(3): 461–462.
30. Yang Y, Wang Y, Zhou Y, et al. Validity of the functional gait assessment in patients with Parkinson disease: construct, concurrent, and predictive validity. *Phys Ther* 2014; 94(3): 392–400.
31. Brusse KJ, Zimdars S, Zalewski KR, et al. Testing functional performance in people with Parkinson disease. *Phys Ther* 2005; 85(2): 134–141.
32. Jacobs JV, Horak FB, Tran VK, et al. Multiple balance tests improve the assessment of postural stability in subjects with Parkinson's disease. *J Neurol Neurosurg Psychiatry* 2006; 77(3): 322–326.
33. Jenkins ME, Johnson AM, Holmes JD, et al. Predictive validity of the UPDRS postural stability score and the Functional Reach Test, when compared with ecologically valid reaching tasks. *Parkinsonism Relat Disord* 2010; 16(6): 409–411.
34. Schenkman M, Cutson TM, Kuchibhatla M, et al. Reliability of impairment and physical performance measures for persons with Parkinson's disease. *Phys Ther* 1997; 77(1): 19–27.
35. Bello-Haas VD, Klassen L, Sheppard S, et al. Psychometric properties of activity, self-efficacy, and quality-of-life measures in individuals with parkinson disease. *Physiother Can* 2011; 63(1): 47–57.
36. Da Silva BA, Faria C, Santos MP, et al. Assessing timed up and go in Parkinson's disease: reliability and validity of timed up and go assessment of biomechanical strategies. *J Rehabil Med* 2017; 49(9): 723–731.
37. Foreman KB, Addison O, Kim HS, et al. Testing balance and fall risk in persons with Parkinson disease, an argument for ecologically valid testing. *Parkinsonism Relat Disord* 2011; 17(3): 166–171.
38. Huang SL, Hsieh CL, Wu RM, et al. Minimal detectable change of the timed "up & go" test and the dynamic gait index in people with parkinson disease. *Phys Ther* 2011; 91(1): 114–121.
39. Morris S, Morris ME and Iansek R. Reliability of measurements obtained with the timed "up & go" test in people with Parkinson disease. *Phys Ther* 2001; 81(2): 810–818.
40. Paul SS, Canning CG, Sherrington C, et al. Reproducibility of measures of leg muscle power, leg muscle strength, postural sway and mobility in people with Parkinson's disease. *Gait Posture* 2012; 36(3): 639–642.
41. van Lummel RC, Walgaard S, Hobert MA, et al. Intra-rater, inter-rater and test-retest reliability of an instrumented Timed Up and Go (iTUG) Test in patients with Parkinson's disease. *PLoS ONE* 2016; 11(3): e0151881.
42. Richards M, Marder K, Cote L, et al. Interrater reliability of the unified Parkinson's Disease Rating Scale motor examination. *Mov Disord* 1994; 9(1): 89–91.
43. Stocchi F, Radicati FG, Chaudhuri KR, et al. The Parkinson's Disease Composite Scale: results of the first validation study. *Eur J Neurol* 2018; 25(3): 503–511.
44. Franchignoni F, Giordano A, Ronconi G, et al. Rasch validation of the Activities-specific Balance Confidence Scale and its short versions in patients with Parkinson's disease. *J Rehabil Med* 2014; 46(6): 532–539.
45. Jacobs JV, Horak FB, Van Tran K, et al. An alternative clinical postural stability test for patients with Parkinson's disease. *J Neurol* 2006; 253(11): 1404–1413.
46. Peretz C, Herman T, Hausdorff JM, et al. Assessing fear of falling: can a short version of the Activities-specific Balance Confidence scale be useful? *Mov Disord* 2006; 21(12): 2101–2105.
47. Spagnuolo G, Faria CD, Da Silva BA, et al. Are functional mobility tests responsive to group physical therapy

intervention in individuals with Parkinson's disease.*? Neurorehabilitation* 2018; 42(4): 465–472.
48. Paul SS, Lester ME, Foreman KB, et al. Validity and reliability of two-dimensional motion analysis for quantifying postural deficits in adults with and without neurological impairment. *Anat Rec* 2016; 299(9): 1165–1173.
49. Duncan RP, Leddy AL, Cavanaugh JT, et al. Accuracy of fall prediction in Parkinson disease: six-month and 12-month prospective analyses. *Parkinsons Dis* 2012; 2012: 237673.
50. Leddy AL, Crowner BE and Earhart GM. Functional gait assessment and balance evaluation system test: reliability, validity, sensitivity, and specificity for identifying individuals with Parkinson disease who fall. *Phys Ther* 2011; 91(1): 102–113.
51. Joanne DiFrancisco-Donoghue M-KJ, Therasa Apononski, Werner WilliamG, et al. The reliability of the sensory organization test in Parkinson's disease to identify fall risk. *Int J Neur Phys Ther* 2016; 2(5): 39–43.
52. Jonasson SB, Nilsson MH and Lexell J. Psychometric properties of four fear of falling rating scales in people with Parkinson's disease. *BMC Geriatr* 2014; 14(1): 66.
53. Almeida LRS, Valenca GT, Negreiros NN, et al. Comparison of self-report and performance-based balance measures for predicting recurrent falls in people with Parkinson disease: cohort study. *Phys Ther* 2016; 96(7): 1074–1084.
54. Dibble LE, Christensen J, Ballard DJ, et al. Diagnosis of fall risk in Parkinson disease: an analysis of individual and collective clinical balance test interpretation. *Phys Ther* 2008; 88(3): 323–332.
55. Park J, Koh SB, Kim HJ, et al. Validity and reliability study of the Korean Tinetti Mobility Test for Parkinson's disease. *J Mov Disord* 2018; 11(1): 24–29.
56. Nilsson MH, Drake AM and Hagell P. Assessment of fall-related self-efficacy and activity avoidance in people with Parkinson's disease. *BMC Geriatr* 2010; 10: 78.
57. Jonasson SB, Nilsson MH and Lexell J. Psychometric properties of the original and short versions of the Falls Efficacy Scale-International (FES-I) in people with Parkinson's disease. *Health Qual Life Outcomes* 2017; 15(1): 116.
58. Thomas M, Jankovic J, Suteerawattananon M, et al. Clinical gait and balance scale (GABS): validation and utilization. *J Neurol Sci* 2004; 217(1): 89–99.
59. Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012; 21(4): 651–657.
60. Fitzpatrick R, Davey C, Buxton MJ, et al. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998; 2: i-iv, 1–74.
61. Brenner MH, Curbow B and Legro MW. The proximal-distal continuum of multiple health outcome measures: the case of cataract surgery. *Med Care* 1995; 33: AS236–AS244.
62. Schenkman M, Ellis T, Christiansen C, et al. Profile of functional limitations and task performance among people with early-and middle-stage Parkinson disease. *Phys Ther* 2011; 91(9): 1339–1354.
63. Almeida LR, Valença GT, Negreiros NN, et al. Predictors of recurrent falls in people with Parkinson's disease and proposal for a predictive tool. *J Parkinsons Dis* 2017; 7(2): 313–324.
64. Baggio JA, Curtarelli Mde B, Rodrigues GR, et al. Validation of the Brazilian version of the clinical gait and Balance Scale and comparison with the Berg Balance Scale. *Arq Neuropsiquiatr* 2013; 71(9a): 621–626.
65. Behrman AL, Light KE, Flynn SM, et al. Is the functional reach test useful for identifying falls risk among individuals with Parkinson's disease? *Arch Phys Med Rehabil* 2002; 83(4): 538–542.
66. Brincks J, Callesen J, Johnsen E, et al. A study of the validity of the Six-Spot Step Test in ambulatory people with Parkinson's disease. *Clin Rehabil* 2019; 33: 1206–1213.
67. Browne J, O'Hare N, O'Hare G, et al. Clinical assessment of the quantitative posturography system. *Physiotherapy* 2002; 88(4): 217–223.
68. Candan SA, Çatıker A and Özcan TŞ. Psychometric properties of the Turkish version of the freezing of gait questionnaire for patients with Parkinson's disease. *Neurol Sci Neurophysiology* 2019; 36(1): 44.
69. Chomiak T, Pereira FV and Hu B. The single-leg-stance test in Parkinson's disease. *J Clin Med Res* 2015; 7(3): 182–185.
70. Chong RK, Lee KH, Morgan J, et al. Diagnostic value of the rapid assessment of postural instability in Parkinson's disease (RAPID) questionnaire. *Int J Clin Pract* 2012; 66(7): 718–721.
71. Dibble LE and Lange M. Predicting falls in individuals with Parkinson disease: a reconsideration of clinical balance measures. *J Neurol Phys Ther* 2006; 30(2): 60–67.
72. Giladi N, Shabtai H, Simon E, et al. Construction of freezing of gait questionnaire for patients with Parkinsonism. *Parkinsonism Relat Disord* 2000; 6(3): 165–170.
73. Goetz CG, Tilley BC, Shaftman SR, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord* 2008; 23(15): 2129–2170.
74. Holmes JD, Jenkins ME, Johnson AM, et al. Validity of the Nintendo Wii® balance board for the assessment of standing balance in Parkinson's disease. *Clin Rehabil* 2013; 27(4): 361–366.
75. Kegelmeyer DA, Kloos AD, Thomas KM, et al. Reliability and validity of the Tinetti mobility test for individuals with Parkinson disease. *Phys Ther* 2007; 87(10): 1369–1378.
76. Kleiner AFR, Pacifici I, Vagnini A, et al. Timed up and go evaluation with wearable devices: validation in Parkinson's disease. *J Bodywork Mov Ther* 2018; 22(2): 390–395.

77. Koop MM, Ozinga SJ, Rosenfeldt AB, et al. Quantifying turning behavior and gait in Parkinson's disease using mobile technology. *IBRO Report* 2018; 5: 10–16.
78. Lindholm B, Nilsson MH, Hansson O, et al. External validation of a 3-step falls prediction model in mild Parkinson's disease. *J Neurol* 2016; 263(12): 2462–2469.
79. Mehdizadeh M, Martinez-Martin P, Habibi S-A, et al. Reliability and validity of Fall Efficacy Scale-International in people with Parkinson's disease during on-and off-drug phases. *Parkinsons Dis* 2019; 2019: 6505232.
80. Nilsson MH and Hagell P. Freezing of gait questionnaire: validity and reliability of the Swedish version. *Acta Neurol Scand* 2009; 120(5): 331–334.
81. Parashos SA, Elm J, Boyd JT, et al. Validation of an ambulatory capacity measure in Parkinson disease: a construct derived from the Unified Parkinson's Disease Rating Scale. *J Parkinsons Dis* 2015; 5(1): 67–73.
82. Paul SS, Canning CG, Sherrington C, et al. Three simple clinical tests to accurately predict falls in people with Parkinson's disease. *Mov Disord* 2013; 28(5): 655–662.
83. Petersen C, Steffen T, Paly E, et al. Reliability and minimal detectable change for sit-to-stand tests and the functional gait assessment for individuals with Parkinson disease. *J Geriatr Phys Ther* 2017; 40(4): 223–226.
84. Ross A, Yarnall AJ, Rochester L, et al. A novel approach to falls classification in Parkinson's disease: development of the Fall-Related Activity Classification (FRAC). *Physiotherapy* 2017; 103(4): 459–464.
85. Taghizadeh G, Martinez-Martin P, Fereshtehnejad S-M, et al. Psychometric properties of the Berg Balance Scale in idiopathic Parkinson'disease in the drug off-phase. *Neurol Sci* 2018; 39(12): 2175–2181.
86. Visser M, Marinus J, Bloem BR, et al. Clinical tests for the evaluation of postural instability in patients with Parkinson's disease. *Arch Phys Med Rehabil* 2003; 84(11): 1669–1674.
87. Yang Y, Wang Y, Zhou Y, et al. Reliability of functional gait assessment in patients with Parkinson disease: interrater and intrarater reliability and internal consistency. *Medicine* 2016; 95(34): e4545.