

The Reliability of Balance, Gait, and Muscle Strength Test for the Elderly with Dementia: A Systematic Review

Han-Suk Lee, PT, PhD · Sun-Wook Park, PT, PhD[†]

Dept. of Physical Therapy, Eulji University

¹Dept. of Physical Therapy, Samsung Medical Center

Received: July 13, 2017 / Revised: July 16, 2017 / Accepted: July 28, 2017

© 2017 J Korean Soc Phys Med

| Abstract |

PURPOSE: To summarize the evaluation tools of balance [Berg Balance Scale (BBS), timed up and Go (TUG), forward reaching test (FRT)], gait [6 m walking Test (6MWT)], and strength [Chair Stand Test (CST)] for patients with dementia.

METHODS: The following databases were searched: Pub MED, Cochrane, Sciences Direct, and Web of Sciences. The inclusion criteria were as follows: 1) repeated measurement design, 2) subjects with dementia, 3) use of testing tools such as the BBS, TUG, FRT, 6MWT, and CST, 4) report the reliability. One reviewer performed the quality assessment of diagnostic accuracy study and two evaluators performed data extraction independently.

RESULTS: Six articles and one letter were included. The interrater reliability of 6MWT, TUG, and CST, were acceptable ($ICC > .90$). However, FRT had unacceptable reliability. In test-retest reliability, only BBS has acceptable reliability ($ICC > .90$). Others had various reliabilities. The

risk of interrater reliability bias was low in all studies. However, the risk of bias of intrarater reliability was low in five studies and moderate in two studies.

CONCLUSION: The interrater reliability of the 6MWT, TUG, and CST were acceptable. However, in test-retest reliability, only BBS has acceptable reliability. Therefore, we suggest the use of BBS to test the balance of dementia patients. In addition, the study of tool reliability according to the subtype of dementia is needed in the future.

Key Words: Dementia, Reliability, Systematic review

I. Introduction

Dementia brings about not only the downgrade of cognitive function, but also the decrease of performance capability, such as balance, muscular strength, and movement capability (Rydwik et al., 2004; Leandri et al., 2009; Kido et al., 2010; Lee and Park, 2016). Patients with dementia would be facing hard time with gait, balance, and mobility (Van Doorn et al., 2003; Feldman et al., 2005; Munoz et al., 2010; Lee and Park, 2017). Demented elders are twice more likely to fall on the ground than those elders whom do not have dementia (Tinetti et al., 1995). Therefore, the study about gait and balance is important.

[†] Corresponding Author : assasun@hanmail.net

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

A recent study showed that there was a positive effect of activity of daily living, functional activity, and fitness on the elders with cognitive disorder when their physical ability was evaluated (Heyn et al., 2004; Littbrand et al., 2009; Ahlskog et al., 2011; Hauer et al., 2012). It is so crucial that useful and proper assessment of performance competence has to serve to the patients with cognitive disorder (Blankevoort et al., 2013). Most frequency tool for elderly person in South Korea are BBS, TUG, FRT for assessing balance; 6MWT, for evaluating gait ability; sit to stand for evaluating strength.

Reliability measurements indicate the degree to which scores of a clinical test are free from measurement errors (Ries et al., 2009). Inter-rater reliability is the degree of agreement among raters and test-retest reliability is the variation in measurements taken by a single person or instrument on the same item, under the same conditions, and in a short period of time. It is recommended to measure both inter-rater reliability and test-retest reliability for accurate reliability of patient with dementia. Patients with dementia might not be able to perform a certain protocol nor to follow any verbal or written instructions due to dealing with lack of understanding skill and provoked anxiety and confusion about the new or unexperienced operations. This is why there are quite many cases of significant decrease in the reliability of assessment. Moreover, many present literatures of the assessment reliability of general movement intervention on the patients with dementia have not satisfyingly been clear with their results nor proposed absolute reliability (Suttanon et al., 2011; Fox et al., 2014). Additionally, the ability of performance may different according to subtype of dementia because there are the communication difference of vascular dementia and Alzheimer disease (Kim and Shin, 2014).

Therefore, it is necessary to figure out such evaluation tool would show convincing reliability on the patients with dementia and compare to type of dementia.

In this study, we summarized evaluation tools of gait

(6MWT) and balance ability (BBS, TUG, FRT) and strength (CST) for the dementia patients and will see which evaluation tool brings the highest reliability. Besides, just for the systemic reviews of literature, this study selected the following hypothesis: there would be a difference among the reliability of each analysis tool of the gait and balance capability depending on the kind of dementia.

II. Methods

The methodology of this study follows a systematic review process originate in evidence-based systematic review of the PRISMA statement (Liberati et al., 2009).

1. Types of studies

Two evaluators were participated the whole procedures independently. They retrieved all relevant literatures in articles, full text according to inclusion criteria. This review study targeted any study to analyze the reliability on assessment tools for evaluating balance, gait and strength. Design of study was repeated measures.

2. Types of participants and intervention

Subjects were elderly people with dementia that was diagnosed by doctor except mild cognitive disorder.

3. Types of outcome measures

Primary outcome measure was a reliability including test-retest reliability or interrater reliability of BBS, TUG, FRT, 6MWT, and CST.

4. Review criteria

Two graduate students did the searching of database separately. Each study evaluated by two authors for the inclusion criteria and discussed any problem each other.

5. Adverse events or side effects

Adverse effects and complications were assessed depending on the description of the authors of each article.

6. Search methods for study identification

This study work through the literature bearing on following sources

Pub MED from 2006- www.ncbi.nlm.nih.gov/sites/entrez?db=PubMed

Cochrane from 2006 - www.thecochranelibrary.com/view/0/index.html

Science Direct from 2006- <http://www.sciencedirect.com/>

Web of Science, from 2006- <https://www.webofknowledge.com>

The period of this work was from January 2006 through December 2016.

7. Search strategy

The search terms include “dementia”, “reliability”, “BBS”, “TUG”, “FRT”, “6meter gait”, “sit to stand”, “balance”, “gait”, and “strength”. All works integrated after two evaluators performed the searching independently in same searching strategy. Any disagreements between evaluators were discussed each other or resolved the problems by third one.

8. Quality assessment

Criteria for quality assessment of articles were cited from the Quality Assessment of Diagnostic Accuracy Studies (Whiting et al., 2011), and checklists used in systematic reviews of diagnostic reliability (Schrama et al., 2014). We performed a list of 11 quality criteria (Table 1).

III. Results

1. Study selection

338 articles were selected after searching with key word. After first screening, 48 articles were selected. Excluded articles have the following reasons that were improper participants (including non-dementia patient) and subjects (not reliability). After getting rid of duplicated articles, 15 studies were chosen. Finally, six articles and one letter that are related to the reliability of BBS, TUG, FRT, 6MWT, and CST that selected among 15 articles. All articles were written in English.

Table 1. Criterion for assessing methodological quality

Items	Answer
1. Were examiner(s) blinded to their own prior test results	Yes, no, unclear
2. Were participants blinded to their own prior test results	Yes, no, unclear
3. Was the order in which subjects/movement/side/joint/device were examined varied?	Yes, no, unclear
4. Was there an appropriate time interval between the measurements to be reasonably sure that the participants' strength could not have been increased or decreased?	Yes, no, unclear
5. Was there an appropriate time interval between the measurements within a day or between days to be reasonably sure that the examiner's characteristics were stable?	Yes, no, unclear
6. Can nonrandom loss to follow-up be ruled out?	Yes, no, unclear
7. Was a representative sample of participations used?	Yes, no, unclear
8. Is replication of the assessment procedure possible?	Yes, no, unclear
9. Was a representative sample of examiners used?	Yes, no, unclear
10. Were the same clinical data available when test results were interpreted as would be available when the rest is used in practice?	Yes, no, unclear
11. Were appropriated statistical measures of intraexaminer reliability used?	Yes, no, unclear

2. Study characteristics

The characteristics of included articles were Table 2. Only one study analyzed the Interrater reliability. Six studies were investigated Test-retest reliability. In five studies, physical therapists were testers. In others studies,

bachelor degree and master degree students, researchers were tester. Sample size of studies was from 12 to 58. In two studies, subjects were AD. In other studies, subjects were dementia.

Table 2. The characteristics of included articles

Study	No. of Patients Mean age (years) Type of dementia	MMSE (SD)	Procedure	Model of ICC	No. of rater Examiners	Type of reliability
Telenius et al. (2016)	33 82 Dementia	15.8	Standardization: yes Unit measurement: yes Single trial used in analysis Mean of 3 trials used in analysis Time for whole test: 30min	2,1	2 Experience:? Profession: PT Training: testing 120 patient in a study 3 months early	Interrater
Muir-Hunter et al. (2015)	15 80.20 (5.03) Dementia	20.0 (5.5)	Standardization: yes Unit measurement: yes Single trial used in analysis: test-retest Between session: 1wk Within session: ? Time for test:?	2.1	2 Experience: assess and treat older adults with balance problems Profession: PT Training: yes	Test-retest Interrater
Suttanon et al. (2011)	14 79.57 (6.19) AD	21.43 (5.00)	Standardization: yes Unit measurement: yes Single trial used in analysis Between session: 1wk Within session: ? Time for test: ?	3.1	1 Experience: six years in clinic Profession: PT Training: ?	Test-retest
Blankevoort et al. (2013)	58 82.47(5.31) Dementia	22.77 (2.13)	Standardization: yes Unit measurement: yes Single trial used in analysis Between session:1wk Within session: ? Time for test:?	2.1	5 Experience: Profession: 5 bachelor degree and master degree Training: yes	Test-retest
Ries et al. (2009)	51 80.71 (8.77) AD	13.1 (8.2)	Standardization: yes Unit measurement: yes Single trial used in analysis : 6MWT Mean score used in analysis: TUG Between session: 30 to 60min (same day) Within session: Time for test:	2.1 (6MWT) 2.2 (TUG)	1 Experience: Profession: PT Training:	Test-retest
Fox et al. (2014)	12 83.25 (9.94) Dementia	?	Standardization: yes Unit measurement: yes Single trial used in analysis Between session:1wk Within session: Time for test:	2.1	1 Experience: measure functional capacity in older adults Profession: Researcher administrator Training: train in the measurement by exercise physiologist	Test-retest

Study	No. of Patients Mean age (years) Type of dementia	MMSE (SD)	Procedure	Model of ICC	No. of rater Examiners	Type of reliability
van Iersel et al. (2007)	39 78.3 Dementia	19.1 (5.2)	Standardization: yes Unit measurement: yes Unclear trial used in analysis Between session:2wk Within session: Time for test:	?	? Experience: Profession: PT, Researcher Training:	Test-retest

MMSE: Mini-mental State Examination, wk: week, PT: physical therapy, No: Number, AD: Alzheimer Disease, 6MWT: 6 Meter Walking Test, TUG: Timed up and Go

Table 3. The assessment of quality within studies

Study	Internal validity						External validity				
	1	2	3	4	5	6	7	8	9	10	11
Teleninus et al. (2016)	Y	?	N	Y	?	?	Y	Y	Y	Y	Y
Muir-Hunter et al. (2015)	N	N	Y	Y	Y	?	Y	Y	Y	Y	Y
Suttanon et al. (2011)	N	N	Y	N	N	Y	Y	Y	Y	N	N
Blankevoort et al. (2013)	Y	N	N	N	Y	Y	Y	Y	N	Y	N
Ries et al. (2009)	N	Y	Y	Y	N	?	N	Y	N	Y	N
Fox et al. (2014)	N	N	Y	N	Y	?	N	Y	?	Y	N
van Iersel et al. (2007)	?	?	Y	?	?	?	?	Y	?	Y	?

3. Risk of bias within studies

There was the quality of studies in Table 3. Five studies showed a low risk of bias (≥ 5 criteria). Two studies exhibited moderate risk of bias (3 or 4 criteria). Two studies, all criteria for external validity was satisfied. The internal validity was lower than external validity.

4. Interrater Reliability of tools of balance, gait and strength

The interrater reliability of tools of balance, gait and strength is shown at Table 4. A total of 2 studies are chosen. All of them were examined dementia subjects.

ICCs ranged from .720 to .995 for assessing BBS. ICCs of assessing for balance ranged from .720 to .980. ICCs of assessing for the strength were one. ICC of assessing for gait was .970. In this study, ICC of assessing for gait was highest. The interrater reliability of TUG, 6MWT, CST were acceptable ($ICC > .90$). However, FRT were nonacceptable reliability. The quality of all studies was low.

The test-retest reliability of tools of balance, gait and strength is shown at Table 5. A total of six studies were chosen. Four studies were examined subjects with dementia and two studies were examined subject with Alzheimer

Table 4. The interrater reliability of tools of balance, gait and strength

Instrument	Type of dementia	Study	Reliability (95% CI)	Risk of Bias	
Balance	BBS	Dementia	Teleninus	ICC= .995	Low
		Dementia	Muir-Hunter	ICC= .72	Low
	TUG	Dementia	Muir-Hunter	ICC= .98	Low
	FRT	Dementia	Muir-Hunter	ICC= .79	Low
Strength	CST	Dementia	Muir-Hunter	ICC = 1	Low
Gait	6MWT	Dementia	Muir-Hunter	ICC = .97	Low

CI: Confidence Interval, BBS: Berg Balance Scale, TUG: Timed Up and Go, FRT: Forward Reach Test, CST: Chair Stand Test, 6MWT: 6 Meter Walking Test

Table 5. The test-retest reliability of tools of balance, gait and strength

Instrument	Type of dementia	Study	Reliability (95% CI)	Risk of Bias	
Balance	BBS	Dementia	Muir-Hunter	ICC=.95	Low
		Dementia	van Iersel	ICC=.97	Moderate
	TUG	Dementia	Muir-Hunter	ICC=.72	Low
		Dementia	Fox	ICC=.857	Moderate
		Dementia	van Iersel	ICC=.97	Moderate
		Dementia	Blankevoort	ICC=.94	Low
		AD	Suttanon	ICC=.757	Low
		AD	Ries	ICC=.987	Low
	FRT	Dementia	Muir-Hunter	ICC=.81	Low
		Dementia	Fox	ICC=.384	Moderate
		Dementia	Blankevoort	ICC=.84	Low
		AD	Suttanon	ICC=.840	Low
Strength	CST	Dementia	Fox	ICC=.966	Moderate
AD		Suttanon	ICC=.797	Low	
Gait	6MWT	AD	Ries	ICC=.987	Low
		Dementia	Blankevoort	ICC=.86	Low
	2.4MWT	Dementia	Fox	ICC=.676	Moderate

CI: Confidence Interval, ICC: Intraclass Correlation Coefficient, BBS: Berg Balance Scale, TUG: Timed Up and Go, FRT: Forward Reach Test, CST: Chair stand Test, 6MWT: 6 Meter Walking Test

disease. ICCs of assessing for balance ranged from .384 to .987. Only BBS has acceptable reliability (ICC>.90) and highest reliability among balance test. The range of reliability of BBS was from .97 to .95. However, the range of reliability of TUG was variable from .72 to .97. Half of studies show non-acceptable reliability (ICC<.90).

There is a large difference between studies in reliability

of FRT test from .384 to .84. FRT test show lowest reliability among balance test. All studies show non-acceptable reliability (ICC<.90).

ICCs of assessing for strength ranged from .966 to .797. Only one studies show acceptable reliability (ICC>.90). ICCs of assessing for gait ranged from .987 to .676. Only one studies show acceptable reliability (ICC>.90).

When comparing the reliability of two groups (dementia: AD), the reliability of AD group was higher than dementia group in 6metWK. However, the reliability of dementia group was higher than AD group in CST. The reliability of TUG, FRT are almost same in both group. The quality of most studies was low except two studies.

IV. Discussion

This study was included seven studies exploring the reliability of tools measuring balance, gait and strength for subject with dementia. The interrater reliability of TUG, 6MWT and CST were acceptable ($ICC > .90$). However, FRT were nonacceptable reliability. In test-retest reliability, only BBS has acceptable reliability ($ICC > .90$). Others have various reliabilities.

Overall, the quality of studies was good including five studies having low risk of bias, two studies showing moderate risk of bias. Specially, blinding of examiner and subject in data was poor.

Sit to stand movements are essential activity during daily life and leg extension muscles power needed for performing this activity. Testing of leg extension muscles power may reflect the function level in older people and many researchers have used the test for older people (Cheng et al., 2014). However, there are few for studying of subject with dementia. In addition, the comparing of reliability of each study is necessary for selecting proper tool for subject with dementia. Therefore, we investigate the reliability of sit to stand test for testing the leg extension power and found the reliability from $ICC = .966$ to $ICC = .797$. The reliability was different between the characteristics of subjects. The reliability of subject group with dementia is higher than subject group with AD. The reason that two group was different might be the understanding the direct of evaluator. One of main problem of AD is the communication ability. Subject with AD might

have more difficulty to understanding of direction of tester than the subject group with dementia. Therefore, when evaluators assess the subject with AD, they should consider the test-retest reliability.

BBS has developed for evaluate balance in the elderly and has been used to measure balance in patients with various condition. In South Korea, BBS is widely used for elderly people in clinic.

Downs et al. (2013) suggested the BBS has high intra and inter-rater reliability in system review. However, the participants who were included the study have various condition like Parkinson's disease, multiple sclerosis, spinal cord injury, outpatients. In our study, we researched only subject with dementia. We found that BBS has higher reliability among balance tools.

TUG test is common tool in clinic for measuring mobility and fall risk in older adults because it is easy to use in field. Therefore, it is meaningful to identify the reliability of TUG.

The reliability of TUG revealed various depending on researchers from $ICC = .72$ to $ICC = .97$ in this study. In comparing with BBS, TUG test can affected by various factors including age, cognitive status, muscle strength, balance control and so on (Whitney et al., 2005; Chen and Chou, 2017). Therefore, the participants can be affected by various condition between test and retest session.

In Reis study showing highest reliability among studies, they performed the test-retest in same day (Ries et al., 2009). Additionally, they recruited lots of number of participants ($n = 51$). In contrast, the study of Muir-Hunter (Muir-Hunter et al., 2015) showing lowest reliability among studies performed only 15 participants apart from 1 week. Therefore, various factors are eliminated by researches in Reis study. According to Streiner et al. (2015), it is necessary to identify psychometric properties of test, which use in new environment or people with different condition. Therefore, when assessor use TUG test in clinic, they must control the condition of test.

FRT and 6MWT are a simple tool of standing balance and gait ability. Duncan et al. (1990) reported the FRT is validity and reliable tool for elderly people but Rockwood et al. (2000) found the FRT is not available for heterogeneous elderly people. Therefore, even though the test is simply to use, the reliability can be differ from participants or other factor.

In our study, we found the reliability of FRT was various depending of researchers from ICC=.384 to ICC=.84. In addition, the reliability of 6MWT was somewhat high from ICC=.987 to ICC=.86. However, the reliability of 2.4MWT was lower than the reliability of 6MWT even though the purpose of test was same to evaluate gait ability.

We can suspect the reason that might be the ability of tester. In case of Fox (Fox et al., 2014) study, the tester was researcher not physical therapist. He was educated from exercise physiologist before researching. That is why the reliability of FRT and gait test in Fox study is lower than other studies.

This study reviewed the data of relative reliability except absolute reliability and did not use the statistical method like meta-analysis. Therefore, the results of this study have little objectivity.

The small numbers of studies were included in our study because of elaborate inclusion criteria. Therefore, it is difficult to generalize the result.

V. Conclusion

In conclusion, this system review has indicated the interrater reliability of TUG, 6MWT and CST were acceptable (ICC>.90). In test- retest reliability, only BBS has acceptable reliability (ICC>.90). Specially, FRT test revealed the low interrater reliability and test-retest reliability and the difference depending on subtype of dementia. Therefore, we suggest the use of BBS for testing of balance to subject with dementia. Also, the study of

reliability of tool according to subtype of dementia is needed in the future.

Acknowledgements

This study was sponsored by Grant No. EJRG 16-12 from Eulji University in 2016.

References

- Ahlskog JE, Geda YE, Graff-Radford NR, et al. Physical exercise as a preventive or disease-modifying treatment of dementia and brain aging. *Mayo Clin Proc.* Elsevier. 2011;86(9):876-84.
- Blankevoort CG, van Heuvelen MJ, Scherder EJ. Reliability of six physical performance tests in older people with dementia. *Phys Ther.* 2013;93(1):69-78.
- Chen TB, Chou LS. Effects of muscle strength and balance control on sit-to-walk and turn durations in the timed up and go test. *Arch Phys Med Rehabil.* 2017. Apr 30. pii: S0003-9993(17)30265-4. doi: 10.1016/j.apmr.2017.04.003.
- Cheng YY, Wei SH, Chen PY, et al. Can sit-to-stand lower limb muscle power predict fall status? *Gait Posture.* 2014;40(3):403-7.
- Downs S, Marquez J, Chiarelli P. The Berg Balance Scale has high intra-and inter-rater reliability but absolute reliability varies across the scale: a systematic review. *J physiother.* 2013;59(2):93-9.
- Duncan PW, Weiner DK, Chandler J, et al. Functional reach: a new clinical measure of balance. *J Gerontol.* 1990;45(6):M192-M7.
- Feldman HH, Van Baelen B, Kavanagh SM, et al. Cognition, function, and caregiving time patterns in patients with mild-to-moderate Alzheimer disease: a 12-month analysis. *Alzheimer Dis Assoc Disord.* 2005;19(1):

- 29-36.
- Fox B, Henwood T, Neville C, et al. Relative and absolute reliability of functional performance measures for adults with dementia living in residential aged care. *Int Psychogeriatr*. 2014;26(10):1659-67.
- Hauer K, Schwenk M, Zieschang T, et al. Physical training improves motor performance in people with dementia: a randomized controlled trial. *J Am Geriatr Soc*. 2012;60(1):8-15.
- Heyn P, Abreu BC, Ottenbacher KJ. The effects of exercise training on elderly persons with cognitive impairment and dementia: a meta-analysis. *Arch Phys Med Rehabil*. 2004;85(10):1694-704.
- Kido T, Tabara Y, Igase M, et al. Postural instability is associated with brain atrophy and cognitive impairment in the elderly: the J-SHIP study. *Dement Geriatr Cogn Disord*. 2010;29(5):379-87.
- Kim YS, Shin MS. The Comprehension of Speech Acts Ability in Alzheimer's disease and vascular dementia. *J Rehabil Psychol*. 2014;21(3):349-72.
- Leandri M, Cammisuli S, Cammarata S, et al. Balance features in Alzheimer's disease and amnesic mild cognitive impairment. *J Alzheimers Dis*. 2009;16(1):113-20.
- Lee HS, Park SW. Assessment of Gait as a Diagnostic Tool for Patients with Dementia. *J Korean Soc Phys Med*. 2017;12(2):129-36.
- Lee HS, Park YJ. The Effect of Physical Activity Program for Elderly with Dementia on Cognitive Function: Meta-Analysis of Studies in Korea. *J Korean Soc Phys Med*. 2016;11(2):115-21.
- Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS medicine*. 2009; 6(7):e1000100.
- Littbrand H, Lundin-Olsson L, Gustafson Y, et al. The effect of a high-intensity functional exercise program on activities of daily living: a randomized controlled trial in residential care facilities. *J Am Geriatr Soc*. 2009;57(10):1741-9.
- Muir-Hunter SW, Graham L, Montero Odasso M. Reliability of the Berg Balance Scale as a Clinical Measure of Balance in Community-Dwelling Older Adults with Mild to Moderate Alzheimer Disease: A Pilot Study. *Physiother Can*. 2015;67(3):255-62.
- Munoz VM, van Kan GA, Cantet C, et al. Gait and balance impairments in Alzheimer disease patients. *Alzheimer Dis Assoc Disord*. 2010;24(1):79-84.
- Ries JD, Echternach JL, Nof L, et al. Test-retest reliability and minimal detectable change scores for the timed "up & go" test, the six-minute walk test, and gait speed in people with Alzheimer disease. *Phys Ther* 2009;89(6):569-79.
- Rockwood K, Awalt E, Carver D, et al. Feasibility and measurement properties of the functional reach and the timed up and go tests in the Canadian study of health and aging. *J Gerontol A Biol Sci Med Sci*. 2000;55(2):M70-3.
- Rydwik E, Frandin K, Akner G. Effects of physical training on physical performance in institutionalised elderly patients (70+) with multiple diagnoses. *Age Ageing*. 2004;33(1):13-23.
- Schrama PP, Stenneberg MS, Lucas C, et al. Intraexaminer reliability of hand-held dynamometry in the upper extremity: a systematic review. *Arch phys Med Rehabil*. 2014;95(12):2444-69.
- Streiner DL, Norman GR, Cairney J. *Health measurement scales: a practical guide to their development and use*. Oxford University Press, USA. 2015.
- Suttanon P, Hill KD, Dodd KJ, et al. Retest reliability of balance and mobility measurements in people with mild to moderate Alzheimer's disease. *Int Psychogeriatr*. 2011;23(7):1152-9.
- Tinetti ME, Doucette J, Claus E, et al. Risk factors for serious injury during falls by older persons in the community. *J Am Geriatr Soc*. 1995;43(11):1214-21.

Van Doorn C, Gruber-Baldini AL, Zimmerman S, et al. Dementia as a risk factor for falls and fall injuries among nursing home residents. *J Am Geriatr Soc.* 2003;51(9):1213-8.

Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic

accuracy studies. *Ann Intern Med.* 2011;155(8):529-36.

Whitney JC, Lord SR, Close JC. Streamlining assessment and intervention in a falls clinic using the Timed Up and Go Test and Physiological Profile Assessments. *Age ageing.* 2005;34(6):567-71.